Origins of Gender Norms:
Sibling Gender Composition and Women's Choice of Occupation and Partner

Anne Ardila Brenøe*
University of Copenhagen and IZA

April 25, 2018

Abstract

I examine how the childhood family environment—more precisely, sibling gender composition—affects women's gender identity, measured through their choice of occupation and partner. Using Danish administrative data, I causally estimate the effect of having a second-born brother relative to a sister for first-born women. The results show that women with a brother acquire more traditional gender norms with negative consequences for labor earnings. I provide evidence of increased gender-specialized parenting in families with mixed sex children, suggesting a stronger transmission of traditional gender norms. Finally, I find indications of persistent effects to the next generation of girls.

JEL classification: I2, J1, J3.
Keywords: Gender identity, sibling gender, occupational choice, family formation.

*Center for Economic Behavior and Inequality (CEBI), Department of Economics, University of Copenhagen, Øster Farimagsgade 5, Building 26, 1353 Copenhagen K, Denmark. Institute for the Study of Labor (IZA), Schaumburg-Lippe-Str. 5-9, 53113 Bonn, Germany. aab@econ.ku.dk. This paper has previously circulated with the title *"Sibling Gender Composition and Participation in STEM Education"*. I thank Marianne Bitler, David Card, Christian Dustmann, Ernst Fehr, Jennifer Graves, Mette Gørtz, Shelly Lundberg, Yana Rodgers, Philip Rosenbaum, Heather Royer, Matti Sarvimäki, Jenna Stearns, Jakob Egholt Søgaard, Melanie Wasserman, Ulf Zölitz, Josef Zweimüller, and seminar participants at the University of Copenhagen (Department of Economics and Department of Sociology), UC-Santa Barbara, UC-Berkeley, UC-Davis, IZA Summer School in Labor Economics 2017, Workshop: Education, Skills, and Labor Market Outcomes 2017, CEN Workshop 2017, IWAEE 2017, Lund University, 2^{nd} IZA Workshop: Economics of Education, Copenhagen Business School, Aarhus University, University of Southern Denmark, DGPE 2017, University of New South Wales, University of Sydney, AASLE 2017, CAM Workshop 2017, ASSA Meetings 2018, University of Vienna, University of Essex, University of Zurich, Hanken School of Economics, DIW Berlin, and Monash University for helpful discussions and comments.

1 Introduction

Across most OECD countries, women today attain more education than men and participate almost equally in the labor force (OECD, 2016; OECD, 2017). But why do women keep choosing fields of study leading to substantially lower-paid occupations (Blau and Kahn, 2016)? Although the barriers to women's participation in education and the labor force have been removed in the attempt to reach gender equality, gender identity still plays an important role for gender differences in behavior and subsequently in economic outcomes (Akerlof and Kranton, 2000; Bertrand, 2011; Goldin, 2014). To really understand why women continue behaving in ways leading to inferior labor market outcomes relative to the ones of men, we need to better understand the origins of—especially women's—gender norms. In this study, I focus on the importance of one key aspect of the childhood family environment—sibling gender composition—for women's socialization and development of gender conformity.

The family constitutes an essential facet of a child's socialization process. Parents act as important role models and transmit gender norms to their children (Farré and Vella, 2007; Fernández et al., 2004; Humlum et al., 2017; Johnston et al., 2013; Kleven et al., 2018). Siblings, at the same time, are close peers during childhood and often sustain long-lasting relationships throughout life (McHale et al., 2013). A child's birth order in the sibship influences, for instance, educational attainment and the development of personality traits through social family interactions (Brenøe and Molitor, 2018; Black et al., 2005, 2017; Lehmann et al., 2016). Sibling gender composition might additionally have a crucial impact on how siblings interact with each other as well as how parents interact with their children (McHale et al., 2003). Parents might, for instance, invest differently in their children depending on the children's gender composition which, in turn, could alter the intergenerational transmission of gender norms.

To examine how sibling gender composition affects the development of women's gender identity, I use high-quality administrative data for the total population in Denmark from 1980 through 2016. With this comprehensive data set, I evaluate women's gender identity through their revealed gender conformity in terms of their choice of occupation and partner from age 31 through 40 (proxied by the gender share in their own and their partner's occupations, respectively). To provide causal estimates of the impact of sibling gender, I exploit the random assignment of the second child's gender in families with a first-born daughter, conditional on the parents having a second child. The crux of my identification strategy is thus to compare the choices for first-born women with a second-born brother to those with a second-born sister. Sibling gender composition has a small impact on family size, yet, I show that family size is not a confounding factor for the effect of sibling gender composition on women's gender conformity. This empirical approach distinguishes itself from previous studies on sibling gender composition, as they generally include all siblings both in the measure of sibling gender composition and in the estimation sample.[1] Considering all siblings is problematic, however, as the final sibling gender composition in a sibship is endogenous. Therefore, studying the effects of older siblings' gender on younger siblings' outcomes may lead to selection bias. If parents, for example, decide to have a second child depending on their first child's gender and if parents with different gender preferences raise their children differently, the estimated effects would be biased. By focusing on the second-born child's gender, I avoid selection bias, as parents do not know the gender of their unborn child when deciding to have another child.

The setting for this study is ideal, as Denmark has been one of the front runners in terms of gender equality for decades. Women from the cohorts of study (1962–1975) attain slightly more education than men[2] and importantly, labor force participation is not gendered. That labor market participation and family formation are not associated with gender identity is a unique (and very essential) feature for the empirical analysis, thereby removing concerns regarding selection into having an observation on choice of occupation, choice of partner, and the outcomes of a first-born child. Yet, pronounced gender differences in occupational choice still persist. Women are, for example, still heavily underrepresented in occupations within Science, Technology, Engineering, and Mathematics (STEM). Therefore, the setting is, in many ways, comparable to the conditions faced by women in other developed countries today.

My results show that having a second-born brother relative to a sister increases first-born women's gender conformity: women with a brother work in more female-dominated occupations during their 30s and choose more traditional partners. In particular, women with a brother are 7.4 percent less likely to work within STEM. In other words, having a brother decreases women's probability of participating in traditionally male-dominated occupations. STEM is one important example of such occupations due to its potential consequences for the individual woman and society, given the higher returns to STEM fields and the need for a talented STEM workforce

[1]E.g. Amin (2009); Anelli and Peri (2014); Bauer and Gang (2001); Butcher and Case (1994); Conley (2000); Cools and Patacchini (2017); Cyron et al. (2017); Hauser and Kuo (1998); Kaestner (1997); Oguzoglu and Ozbeklik (2016); Rao and Chatterjee (2017). The only exceptions from such strategy are Cronqvist et al. (2015) and Peter et al. (2015), investigating the effect of a co-twin's gender on financial risk taking, education, earnings, and family formation. Moreover, Gielen et al. (2016) employ a difference-in-differences strategy to estimate the effect of having a male twin on earnings; yet, their interest is whether exposure to prenatal testosterone (rather than sibling gender composition per se) has an effect on earnings. Cools and Patacchini (2017) and Rao and Chatterjee (2017) both provide a robustness check of their estimates on wages in which they only consider the sex of a next younger sibling.

[2]This is a fortunate feature, as previous studies on sibling gender composition have been concerned with the potential role of differential parental investment in daughters when also having sons, as parents in more traditional societies tend to favor boys. In Denmark, on average, parents do not favor one gender over the other (Andersson et al., 2006). Therefore, sibling gender is not associated with financial constraints.

to sustain long-run economic growth (Altonji et al., 2015; Kirkebøen et al., 2016; Peri et al., 2015). As male-dominated occupations are generally better paid, I further show that women with a brother earn less than those with a sister. I provide evidence that differences in labor market participation and family formation cannot explain the effects on occupational choice or labor earnings. While the main analysis concerns the development of women's gender identity, I also briefly present the results from a similar analysis for men (Section 6). Consistent with the findings for women, the results suggest that having an opposite sex sibling enhances men's gender identity.

The effect of sibling gender on women's gender conformity propagates through life and is already visible when considering their educational choice. While sibling gender has no effect on educational attainment or achievement, women with a brother complete less male-dominated educations. As an example, having a brother decreases women's probability of completing any field-specific STEM education by 11.3 percent. This effect on women's falling out of STEM fields is already present in their first educational choice after compulsory schooling at age 16. The key finding that women with a brother acquire more gender-typed human capital motivates an analysis of whether the effects persist into the human capital formation of the next generation. Remarkably, the results show that daughters' comparative advantage in language over math in school is larger for those with a more gender-conforming mother, i.e. for daughters of mothers with a brother relative to daughters of mothers with a sister. Thus, I find evidence of very persistent long-run consequences of women's childhood family environment.

Why does sibling gender affect the development of women's gender identity? The effect of having a brother could go through either child-parent and/or child-sibling interactions.[3] I provide compelling evidence in favor of the former channel by showing that parents of mixed sex children invest their time more gender-specifically in their first-born child than parents of same sex children. The results from heterogeneity analyses further indicate that the effect of having a brother is largest for women from more traditional families. These findings are consistent with the argument, similar to the one put forward in the same sex education literature (Booth et al., 2013; Schneeweis and Zweimüller, 2012), that having an opposite sex sibling increases girls' exposure to gender-stereotypical behavior and thereby increases their inclination to acquire more traditional gender norms. In support of this argument, Cools and Patacchini (2017) and Rao and Chatterjee (2017) provide some indications that women with brothers hold more traditional gender attitudes than those without brothers.

My focus on the social environment and the origins of gender norms is consonant with recent studies that trace gender gaps in educational outcomes to factors such as

[3]The impact of having a brother on gender identity could also theoretically be due to changes in ability and parental resource constraints. However, I rule this out by showing that sibling gender does not affect school performance or attainment.

teacher stereotypes, the gender of school peers and teachers, and parental and sibling role models.[4] A strand of the literature shows, for instance, that gender-stereotypes in the school environment affect the gender gap in math test scores.[5] Fewer studies, however, trace effects into outcomes with consequences for economic well-being in adulthood, such as field of education, working decisions, and earnings—in part, due to limited data availability. Some exceptions exists, however. For instance, Olivetti et al. (2016) show that having more female peers with working mothers during adolescence increases young women's probability of working and Kleven et al. (2018) show that women's child penalty on wages is largest for those from more traditional families. The literature on sibling gender composition is small and has predominantly been concerned with educational attainment, while a couple of more recent papers focus on wages.[6] The evidence on educational attainment is overall mixed, while studies on wages reach a more consistent finding that both male and female wages are negatively associated with having an opposite sex sibling—similar to my findings.

This paper makes five important contributions to the existing literature. First, I provide a comprehensive analysis of how sibling gender composition causally affects the development of women's gender identity, using two novel measures of gender conformity. Second, the large sample size and administratively reported occupations provide precisely estimated effects on the gender conformity of women's occupational choice.[7] Third, to the best of my knowledge, I am the first to consider the gender conformity of the choice of women's partner, which is again only possible due to the rich data set, as I am able to identify all partners and their occupations without relying on self-reports. Fourth, I document lasting effects to the next generation of girls, thereby stressing the persistence of gender norms. Fifth, I conduct a large quantitative analysis of how sibling gender composition affects child-parent interactions, thereby providing a detailed picture of an important channel through which the effects on gender identity operate.

[4]See e.g. Anelli and Peri (2014, 2016); Bottia et al. (2015); Brenøe and Lundberg (2017); Brenøe and Zölitz (2018); Carrell et al. (2010); Cheng et al. (2017); Zölitz and Feld (2017); Humlum et al. (2017); Joensen and Nielsen (2017); Johnston et al. (2013); Oguzoglu and Ozbeklik (2016).

[5]Several studies find that having a gender-stereotypical teacher increases the math test score gap, mainly by decreasing girls' performance (Alan et al., 2017; Carlana, 2017; Lavy and Sand, 2015; Lavy and Megalokonomou, 2017).

[6]See the references in Footnote 1. A general problem, though, is small sample sizes, often resulting in quite imprecise estimates, and potential biases.

[7]This is in contrast to the only few existing studies that have attempted to consider occupational outcomes, such as an occupational prestige score and binary indicators for occupational groups (Cools and Patacchini, 2017; Rao and Chatterjee, 2017). Their sample sizes ($< 5,000$) have, however, been too small to allow for any clear conclusions; the estimates of the signs are generally consistent with my main findings, though.

2 Empirical Strategy

The aim is to estimate the causal effect of sibling gender composition on the formation of women's gender identity. Simply comparing women from families with different gender compositions would, however, not provide valid estimates of the causal effect of sibling gender composition due to selection. The final gender composition in a family is endogenous, as parents decide whether or not to have more children after each childbirth and thereby when knowing their current children's gender composition. If parents' decision to have a second child depends on the first child's gender and if such gender preferences also affect how parents raise their children, it is not possible to estimate the causal effect of "current" (first-born) children's gender on "future" (second-born) children's outcomes because not all "future" children are born.[8]

To reach the goal of estimating the causal effect of sibling gender composition, I focus on the random assignment of the second-born child's gender. Because parents do not know the gender of a subsequent child when they make the decision to progress to the next parity, I *can* causally estimate the effect of a "future" child's gender on "current" children's outcomes. Thus, I leverage the random assignment of the second child's gender in families with a first-born daughter, conditional on having a second child. In other words, I compare first-born women who have a second-born brother to first-born women who have a second-born sister. Thereby, the identifying assumption is that conditional on the first child's gender and conditional on having a second child, the sex of the second child is random.

The empirical specification for the main analysis is:

$$Y_i^{First-Born} = \alpha_0 + \alpha_1 Brother_i^{Second-Born} + X_i'\delta + \nu_i, \tag{1}$$

where $Y_i^{First-Born}$ measures woman i's (who is first-born) gender conformity. The estimate of interest is α_1, representing the effect of having a second-born brother. X_i is a vector of fixed effects for birth municipality, year-by-month of birth, spacing in months to the second-born sibling, maternal age at birth, paternal age at birth, maternal level-by-field of education, and paternal level-by-field of education.[9] ν_i is the error term.[10]

As this strategy only relies on the random assignment of the second child's sex, parents can respond to the gender composition of their first two children in terms of subsequent fertility. Consistent with the literature exploiting sibling sex composition as an instrument for family size (e.g. Angrist and Evans (1998)), Appendix Table A1

[8]Appendix A.1 shows the selection bias problem more formally and discusses other reasons for selection bias than parental gender preferences.

[9]If the parent does not have a field-specific education, I use their field of occupation.

[10]I do not cluster the standard errors; however, the results do not change if I do so.

shows that, for the main sample of the analysis (described in Section 3), having two mixed sex children reduces family size by 0.07 children, on average. Therefore, family size might mediate some of the effect of having a second-born brother if family size has an independent impact on gender identity. Existing studies find that family size does not affect educational attainment in Israel or Norway, using twins as an instrument for family size (Angrist et al., 2010; Black et al., 2005). In Appendix A.2.1, I replicate this finding in the Danish context and show that neither does family size affect the different measures of gender conformity. Appendix A.2.2 provides additional tests of the sensitivity of the findings, which further lend support to the conclusion that the results are robust to family size. Based on this wide battery of tests, family size does not seem to be an important confounder of the effect of sibling gender.

3 Data

3.1 Data and Sample Selection

I use Danish administrative data for the total population from 1980 through 2016. One central feature of this data set, compared to most previous studies on sibling gender composition, is that I can link all children to their parents and siblings. Thus, I observe parents' complete fertility history and thereby, correctly measure the sibling gender composition. Furthermore, I have information on parents' date of birth, length, type, and field of education, labor market attachment, and occupation. For the children, I annually observe labor market outcomes, educational enrollment and completion, fertility, and cohabitation and marital status. Finally, I observe the school performance of the children's children.

I restrict the sample to women born between 1962 and 1975 to be able to study the choice of occupation and partner when these women are in their 30s. Moreover, I only include first-born women, who are the first child to both the mother and father; I exclude immigrants;[11] I only consider individuals who have at least one full sibling (same mother and father) born less than four years apart and who survives the first year of life; I exclude families in which either the first or second child is a twin; and finally, I exclude those few women who die before age 40 or do not live in Denmark at any time between age 31 and 40 when the main outcome variables are measured.[12] I refer to this sample of first-born women as the *main sample*.

[11] For first-generation immigrants, I do not necessarily have complete sibling or parental information. Second-generation immigrants would have represented approximately one percent of the sample, reason for which I decided to exclude them to have a more homogeneous sample. However, including second-generation immigrants does not change the results.

[12] Sibling gender composition does not affect attrition due to these restrictions.

Table 1
Descriptive Statistics on Childhood Family Environment for Sample of First-Born Women

Panel A: Statistic by Gender of the Second-Born Sibling	Sister		Brother		t-test
	Mean	SD	Mean	SD	p-value
	(1)	(2)	(3)	(4)	(5)
Predetermined Characteristics					
Spacing (months)	29.9	9.6	30.0	9.6	0.16
Mother's age (years)	22.9	3.6	22.8	3.6	0.21
Father's age (years)	25.7	4.4	25.6	4.4	0.06
Mother's education (years)	10.9	3.2	10.9	3.2	0.62
Father's education (years)	11.8	3.3	11.8	3.3	0.54
Mother has $\geq$ 12 years of education	50.8	50.0	51.2	50.0	0.28
Father has $\geq$ 12 years of education	65.7	47.5	65.8	47.4	0.85
Both parents have $\geq$ 12 years of edu	41.5	49.3	41.8	49.3	0.33
Mother in care or administration	15.6	36.3	15.8	36.4	0.42
Father in STEM	8.2	27.4	8.3	27.6	0.58
Mother in care/adm & Father in STEM	2.4	15.2	2.4	15.3	0.68
Parental Response to Sex Composition					
Number of siblings	1.7	0.9	1.6	0.9	<0.01
Has $\geq$ 2 siblings	39.9	49.0	34.6	47.6	<0.01
Has $\geq$ 3 siblings	8.4	27.8	7.1	25.6	<0.01
Lives with both bio parents	81.0	39.2	81.1	39.1	0.62
Lives w mother, sib w father	4.6	20.9	9.9	29.9	<0.01
Parents Equal Division of Labor	33.7	47.3	33.4	47.2	0.38
Observations	50,757		52,776		
Panel B: Balancing Test					
Joint F-statistic			0.92		
Prob > F			0.92		

Note: Main sample (first-born women born 1962–1975 with a second-born biological sibling born within four years apart). Panel A shows the average and standard deviation of family background characteristics for first-born women with a second-born sister [Columns (1) and (2)] and brother [Columns (3) and (4)]. Column (5) reports the p-values from t-tests of significance between the averages of the two groups of women. Panel B tests whether the control variables included in X_i in Equation (1) can predict having a second-born brother. F-test of joint significance of all control variables.

Table 1 provides descriptive statistics on the childhood family environment for the the main sample by the gender of the second-born sibling. As expected, these women come from families with similar predetermined family characteristics regardless of sibling gender. On average, spacing to the younger sibling is 2.5 years, mothers are 22.9 years at birth and have 10.9 years of education, while fathers are 25.7 years and have 11.8 years of education. When it comes to characteristics that the parents can manipulate after realizing the gender composition of their first two children, we see that those with two daughters are more likely to have more children, as discussed in Section 2. Meanwhile, the probability of having both parents working equally[13] during childhood or living with both biological parents at age 17 does not differ by sibling gender composition. Among those not living with both parents at age 17, however, we see a clear difference in the family living arrangement: divorced parents with mixed sex children are more likely to live with their same sex child only.

To provide support for the identifying assumption, that sibling gender is random, Column (5) in Panel A tests whether the background characteristics differ by gender of the second-born sibling. Considering the predetermined characteristics, only father's age at birth differ marginally between the two groups.[14] Panel B shows statistics from a balancing test, testing whether the demographic characteristics included in X_i in equation (1) can predict sibling gender. More precisely, it reports the F-test of joint significance of all the covariates in a regression where the outcome is an indicator for having a second-born brother. The F-test strongly rejects joint significance. Thus, this balancing test supports the identifying assumption that the younger sibling's gender is random conditional on the first child's gender and conditional on having a second child.[15]

3.2 Outcome Variables

The three main outcome variables evaluate the degree of women's gender conformity. The first outcome reflects how gender-typed the individual woman's occupational choice is. More precisely, I construct this variable as the natural logarithm of the average male

[13]I define this as the tertile of families in which the parents' division of labor until the child turns 19 years is most equal. More precisely, fathers in this group work at most 62 percent of total parental labor supply. I observe parents' labor supply through a mandated pension scheme (ATP), in which employers contribute for each employee based the number of hours worked.

[14]To account for this small baseline difference, I flexibly control for parental age among a wide range of other fixed effects in the analysis.

[15]The graphs in Appendix Figure A1 illustrate the estimates from an event study of the effect of having a second-born son on a variety of parental socio-economic characteristics. The gender composition of children does not affect parental cohabitation, marital status, length of education, parental employment, or parental annual labor earnings before or around the birth of their first child.

share in the woman's 4-digit occupation codes observed between age 31 and 40.[16] The second outcome measures the share of years between age 31 and 40 the woman works in a high-skilled STEM occupation. The third outcome quantifies how traditional the woman's choice of partner is. This variable measures the natural logarithm of the female share in the partner's occupation.[17] Table 2 provides descriptive statistics on the outcome variables for the main sample of women by sibling gender and for a sample of men that is selected similarly to the main sample, for comparison. We observe a strong degree of gender segregation in occupational choice. While women, on average, have 33 percent men in their occupation, this number is 72 percent for men. Similarly, women's partners have, on average, 28 percent women in their occupation compared to 66 percent for men's partners. Moreover, men are three times more likely than women to work within STEM.

To study potential causes and consequences of occupational choice, I further consider educational and labor market outcomes. I examine labor market outcomes from age 18 through 40 in terms of the labor earnings percentile by age and cohort, work experience, and unemployment history. The earnings percentile provides a standardized measure of relative income that includes individuals with zero earnings, is comparable across cohorts and ages, and is constructed based on the total population. At age 40, women have an average earnings percentile of 49, corresponding to a mean labor income of 320,000 DKK (43,000 EUR). While women only earn 70 percent of men, men and women participate almost equally in the labor market: by age 40, women (men) have 14 (16) years of work experience and 1.8 (1.2) years of unemployment. Similarly, these cohorts of women and men attain almost equal length of education; by age 30, women have on average completed 13.3 years of education and men have completed 13.2 years. In line with the differences in occupational choice, the male share in the highest completed degree is much lower for women (36 percent) than for men (66 percent) and women are much less likely to enroll in and complete any field-specific STEM education.[18]

Furthermore, I examine whether sibling gender affects family formation through age 41. This aspect of women's life might reflect a certain degree of gender-conformity and might at the same time influence labor market outcomes (Bertrand, 2011). First,

[16] I use the Danish version of International Standard Classification of Occupations (DISCO), which I observe from 1991 through 2013.

[17] I define the partner as the mode person with whom the woman cohabits or is married between age 31 and 41. Sibling gender has no impact on women's probability of having an observation on the partner's occupation (not reported). I consider the logarithm of the male share in the woman's own occupation and the logarithm of the female share in her partner's occupation because these measures best approximate a normal distribution rather than considering the logarithm of the male share in both persons' occupations.

[18] See Appendix A.3 for details on the educational outcomes and the educational system in Denmark with emphasis on STEM education.

Table 2
Descriptive Statistics on Outcome Variables for Sample of First-Born Women by Gender of Second-Born Sibling (and First-Born Men for Comparison)

	Women				Men	
	Sister		Brother		Sister or Brother	
	Mean (1)	SD (2)	Mean (3)	SD (4)	Mean (5)	SD (6)
Choice of Occupation and Partner						
Male share in own occupation	33.6	21.1	33.2	20.9	71.6	22.1
STEM occupation	5.2	19.0	4.8	18.1	14.2	30.3
Female share in partner's occ	28.4	21.4	28.0	21.4	66.4	20.3
Labor Market Outcomes at age 40						
Earnings Percentile	49.1	24.8	48.7	24.7	64.4	27.4
Earnings (1,000 2015-DKK)	320.6	197.6	318.6	197.8	460.7	395.3
Work experience (months)	168.9	63.4	168.6	63.7	192.2	69.0
Unemployment (months)	21.4	25.5	21.5	25.6	14.3	21.7
Education by age 30						
Male share in education	36.0	21.5	35.7	21.5	66.4	25.2
Length of education (months)	159.6	26.7	159.5	26.6	158.8	27.4
Academic high school GPA	0.02	0.99	0.01	0.99	0.09	1.03
Any STEM enrollment	8.3	27.6	7.7	26.7	41.6	49.3
Any STEM completion	5.1	21.9	4.5	20.8	30.3	45.9
Marital and Fertility History by age 41						
Cohabit share age 18–41	26.8	21.0	26.0	20.7	23.8	19.6
Married share age 18–41	39.0	27.6	38.9	27.7	30.1	25.5
Has any children	88.7	31.7	88.5	31.9	79.5	40.4
Number of Children	2.0	1.1	2.0	1.1	1.7	1.1
Age at first childbirth	27.3	4.7	27.3	4.7	29.3	4.6
First-Born Child's Grade 9 GPA (standardized with mean 0, SD 1)						
Daughter language	0.37	0.93	0.40	0.92	0.34	0.94
Daughter math	0.13	0.95	0.14	0.96	0.08	0.97
Son language	-0.07	0.96	-0.07	0.97	-0.11	0.97
Son math	0.23	0.95	0.24	0.95	0.19	0.97
Observations	50,757		53,012		108,366	

Note: Main sample (first-born women born 1962–1975 with a second-born biological sibling born within four years apart); the sample of men corresponds to the one of women with the exact same sample selection criteria. Columns (1) and (3) show the average outcome variables for first-born women with a second-born sister and brother, respectively, while Column (5) shows the average for first-born men regardless of the second-born's gender.

I consider the share of years between age 18 and 41 during which the woman cohabits without being married (henceforth *cohabit*) and is married, respectively. Second, I consider the probability of having any children, the number of children, and age at first childbirth conditional on having any children. Although having a partner (and being married) and having children might reflect a greater degree of gender-stereotypical behavior, it is not inevitably the case (Bertrand et al., 2016). Cohabitation could instead reflect non-traditional behavior, as marriage is the tradition. Moreover, the vast majority (89 percent) of women have at least one child and most of those having children have exactly two. Therefore, gender identity may not necessarily influence family formation.

Finally, the last group of outcomes concerns the school performance of the next generation. For this, I consider the outcomes of the first-born child and split the sample by the child's gender.[19] I examine the externally-graded grade point average (GPA) from the grade 9 written language (Danish) and math exams. Both measures are standardized with mean zero and standard deviation (SD) one by exam year for the entire student population. Generally, and as seen from the data, girls perform much better (0.45 SD) than boys in languages, while boys perform slightly better (0.10 SD) than girls in math. Therefore, languages may be perceived as more feminine and math more masculine. Thus, if mothers' gender identity transmits to their children (daughters), we might observe a widening in the gap between language and math performance. Given previous findings, suggesting that mothers influence their daughters more than sons and vice versa for fathers (Brenøe and Lundberg, 2017; Brenøe and Epper, 2018; Humlum et al., 2017), we would mainly expect to observe an effect of the gender of the mother's sibling on daughters' and not on sons' performance.

4 Results

4.1 Gender Identity: Choice of Occupation and Partner

Table 3 shows the main results on the impact of sibling gender on women's choice of occupation and partner, with different control versions. The models in Column (1) show the raw means between first-born women with a second-born sister and those with a second-born brother, while Column (2) includes basic demographic controls. Column (3), the preferred model, further controls for parental education. Finally, Column (4) flexibly adds controls for family size and the sex of potential third- and

[19] Given child gender is independent of the gender of the mother's sibling, this split does not create any bias. Yet, sibling gender might affect the mother's gender preference for her own children and thereby her subsequent fertility choices. Therefore, I only consider women's first-born children. I do not observe any selection into having an observation on a first-born child's outcomes.

Table 3

Effect of Sibling Gender on Choice of Occupation and Partner

	(1)	(2)	(3)	(4)
Panel A: Log(Male Share in Own Occupation)				
Second-Born	-1.16**	-1.17**	-1.22***	-1.29***
Brother	(0.48)	(0.47)	(0.47)	(0.47)
Observations	103,769	103,769	103,769	103,769
Panel B: Share of Years Working in STEM Occupation				
Second-Born	-0.38***	-0.37***	-0.38***	-0.42***
Brother	(0.12)	(0.11)	(0.11)	(0.11)
Observations	103,769	103,769	103,769	103,769
Panel C: Log(Female Share in Partner's Occupation)				
Second-Born	-1.98***	-1.74***	-1.88***	-1.89***
Brother	(0.67)	(0.67)	(0.66)	(0.66)
Observations	95,087	95,087	95,087	95,087
No controls	✓			
Basic controls		✓	✓	✓
Parental education			✓	✓
Family size				✓

All estimates are multiplied by 100 to express effects in percentage/log-points. Standard errors in parentheses; * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Main sample (first-born women born 1962–1975 with a second-born biological sibling born within four years apart). Each Panel-Column presents estimates from separate regressions. *Basic controls* include fixed effects for birth municipality, year-by-month of birth, spacing in months to younger sibling, maternal age at birth, and paternal age at birth. For the own occupation outcomes, basic controls also include dummies for the number of years observed in the income registry from age 31–40 and the number of years observed with a valid occupation code from age 31–40. For partner's occupation, basic controls also include dummies for the partner's number of occupational observations and age at first and last observation. *Parental education* controls include fixed effects for maternal level-by-field of education and paternal level-by-field of education. *Family size* controls include dummies for the number of biological siblings and dummies for the number of children the mother and father potentially have, respectively, from later relationships, and the gender of the third- and fourth-born siblings. The occupational outcomes of the first-born women are measured as mean from age 31–40. The occupational outcome of the partner is measured mainly at ages 31–45 for the partner with whom the woman lived most years from age 31–41.

fourth-born siblings.[20] As family size is an outcome of sibling gender composition, the latter control version might bias the estimates. This control version, however, works as a robustness check of the results, as family size might also be considered a confounding variable. Regardless of the covariates included, the estimates across the different control versions are almost identical, supporting the assumption that sibling gender is random and illustrating that family size is not a principal mediator of the effect of sibling gender (as discussed in more detail in Appendix A.2). Therefore, the rest of this paper proceeds by presenting the results, using the preferred control version in Column (3).

Overall, the results show that having a second-born brother relative to a sister enhances women's gender identity. First-born women with a second-born brother work in occupations with 1.22 percent fewer men compared to first-born women with a second-born sister. Note that this difference in occupational choice is observed well into these women's labor market careers during their thirties (as an average from age 31 through age 40). Consistent with this, having a brother also reduces women's probability of working within STEM by 0.38 percentage points, corresponding to a decrease by 7.35 percent relative to the mean for women with a sister. Consequently, the results clearly show that having a brother induces women to exhibit more traditional choices of occupation. In other words, they are less prone to opt into traditionally male-dominated occupations, of which STEM is one important example.

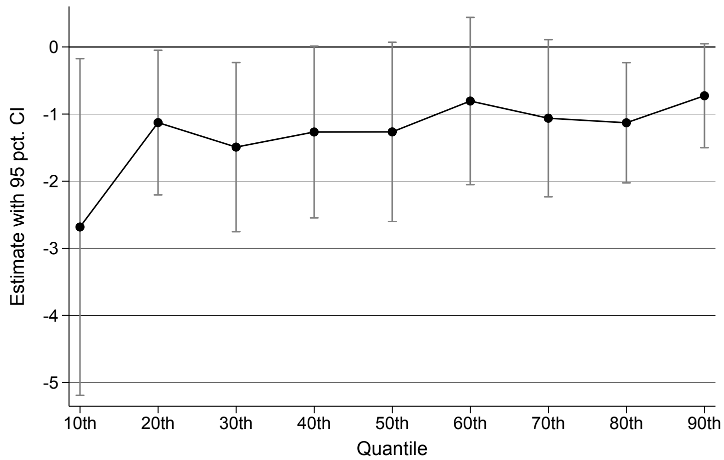
Moreover, sibling gender has a significant impact on the choice of partner in terms of the degree of how gender-typed his occupation is. Having a brother rather than a sister induces women to choose a partner who works in more male-dominated occupations. On average, women with a brother have a partner working in occupations with 1.88 percent fewer women than women with a sister. Not reported, having a brother increases the difference in the male share between the woman's own and her partner's occupation by 0.80 percentage points. These results hereby demonstrate a powerful effect of having a brother on women's choice of gender-stereotypical occupations and partners.

Figure 1 considers whether the effects differ across the different parts of the distribution, by presenting the results from quantile regressions. Both for the male share in the woman's own occupation and the female share in the partner's occupation, the estimates are not statistically significantly different from each other from the tenth through the ninetieth percentiles. Yet, for both measures, the estimates indicate largest effects at the lower part of the distributions. At the tenth percentile, the estimated effects of having a brother relative to a sister are approximately twice the magnitude of the ones seen in Table 3. This suggests that those women who are affected the most by having

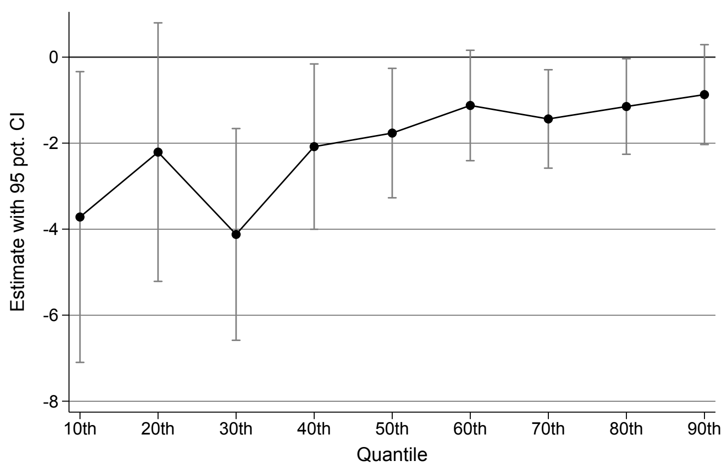
[20] The estimates are identical when not controlling for third- and fourth-born siblings' gender.

Figure 1

Distributional Effects of Sibling Gender on Choice of Occupation and Partner



(a) Log(Male Share in Own Occupation)



(b) Log(Female Share in Partner's Occupation)

Note: All estimates are multiplied by 100 to express effects in log-points. The whiskers represent the 95 percent confidence interval. Main sample (first-born women born 1962–1975 with a second-born biological sibling born within four years apart). All estimates come from separate quantile regressions. All models control for quadratic spacing to the second-born sibling, mother's and father's cubed age at birth, and absorb fixed effects for year of birth, indicators for missing parental age information, and a constant. The models in Graph (a) further control for dummies indicating the number of occupational observations and the models in Graph (b) control for the partner's number of occupational observations and age at first and last observation.

a brother are those who are more traditional than the average.

If the effect of sibling gender, at least partly, goes through the way in which parents treat their children, we might observe some heterogeneity in the effect of having a

Table 4
Heterogeneity: Choice of Occupation and Partner

	Log(Male Share in Own Occupation)	Share of Years in STEM Occupation	Log(Female Share in Partner's Occupation)
	(1)	(2)	(3)
Panel A: Parental Division of Labor During Childhood			
Second-Born	-1.63***	-0.36**	-1.96**
Brother (SBB)	(0.59)	(0.14)	(0.82)
SBB $\times$Equal	1.74*	-0.03	0.09
	(1.01)	(0.25)	(1.42)
Observations	100,020	100,020	91,706
Panel B: Parental Field of Academic Education			
Second-Born	-0.64	-0.19	-1.74**
Brother	(0.54)	(0.13)	(0.75)
SBB$\times$Mother Care/Adm	-1.46	-1.02***	-0.90
	(1.41)	(0.35)	(1.99)
SBB$\times$Father STEM	-3.79*	-0.75	-1.19
	(2.04)	(0.50)	(2.87)
SBB$\times$Mother Care/Adm	2.01	1.09	1.60
$\times$Father STEM	(3.91)	(0.96)	(5.54)
Observations	100,772	100,772	92,406
Panel C: Parental Years of Education			
Second-Born	0.84	-0.21	-1.64
Brother	(0.96)	(0.24)	(1.35)
SBB$\times$Mother≥ 12&Father< 12	-3.06*	-0.35	4.05
	(1.84)	(0.45)	(2.58)
SBB$\times$Mother< 12&Father≥ 12	-2.95**	-0.07	0.64
	(1.36)	(0.33)	(1.91)
SBB$\times$Mother≥ 12&Father≥ 12	-2.32*	-0.29	-1.99
	(1.21)	(0.30)	(1.70)
Observations	100,772	100,772	92,406

All estimates are multiplied by 100 to express effects in percentage/log-points. Standard errors in parentheses; * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Main sample (first-born women born 1962–1975 with a second-born biological sibling born within four years apart). Each Panel-Column presents estimates from separate regressions. All models absorb fixed effects for birth municipality, year-by-month of birth, spacing in months to younger sibling, maternal age at birth, paternal age at birth, maternal level-by-field of education, and paternal level-by-field of education. Columns (1) and (2) also include dummies for the number of years observed in the income registry from age 31–40 and the number of years observed with a valid occupation code from age 31–40. Column (3) also includes dummies for the partner's number of occupational observations and age at first and last observation. The occupational outcomes of the first-born women are measured as mean from age 31–40. The occupational outcome of the partner is measured mainly at ages 31–45 for the partner with whom the woman lived most years from age 31–41.

brother by parental characteristics.[21] Panel A in Table 4 includes an interaction term between sibling gender and an indicator for having parents working (close to) equally during childhood. Remarkably, the effect of having a brother on occupational choice disappears for women coming from more gender-equal families. This suggests that women with more gender-stereotypical parents drive the effect of sibling gender on choosing more female-dominated occupations. Moreover, the results in Panel B suggest that the effect of having a brother is largest for those women with more traditional parents in terms of their educational field. The effects seem to be largest in magnitude for those with a mother who has an academic education within care or administration and for those with a father who has an academic education within STEM.

The effect of having a brother is, furthermore, the largest for those with at least one highly educated parent ($\geq$ 12 years of education) for occupational choice. A highly educated parent will in most cases also imply having a parent with human capital that is traditionally associated with his or her own gender. For instance, most mothers with long education are within care and administration (e.g. nurse, secretary, and office work) and most fathers are within STEM. Therefore, these results again support the previous findings that the effect of having a brother is largest for those with more gender-stereotypical parents. Notably, the results also show that women with both parents having less education do not experience an effect of sibling gender. This suggests that the effect is not due to resource constraints, which has been put forward as a relevant mechanism in the sibling gender composition literature on educational attainment (Amin, 2009; Butcher and Case, 1994). Although the estimates are more imprecisely estimated for the other two outcomes, they are qualitatively consistent with the findings for the male share in the woman's occupation.

Expanding the sample to include individuals spaced up to eight years from their second-born sibling shows that sibling gender does not have an impact for those with long spacing to their sibling [Appendix Figure A2]. Though, the estimated effects by spacing are not statistically significantly different from each other, probably due to the small fraction of children with long spacing to their second-born sibling. This finding that individuals with long spacing to their younger sibling do not experience an effect of sibling gender might indicate the importance of sibling interactions. However, it could also be because parents with children spaced far apart treat the first-born child similarly regardless of the younger sibling's gender.

In sum, these heterogeneities indicate that the effect of having a brother is largest for women from more traditional families. This, in turn, suggests that differences in child-parent interactions are important for the effects of sibling gender composition on the formation of women's gender identity. Ceteris paribus, we would expect that parents

[21] As seen in Table 1, these parental characteristics do not differ by sibling gender composition.

with more gender-stereotypical human capital would reinforce gender-specialization to a larger extent than those parents with less gender-specific human capital (Humlum et al., 2017). Additionally, we would expect that spending more time with one parent with human capital traditional for one gender than with the other parent with human capital traditional for the other gender would influence the child more in the direction of the former parent's interests than the latter parent's. Therefore, the results are consistent with the hypothesis that parents of mixed sex children invest more time in their same sex child than parents of same sex children; Section 5 elaborates more throughly on this.

4.2 Labor Market Outcomes

As female-dominated occupations typically pay lower wages, an important consequence of the results on occupational choice may be reflected in lower labor earnings. To study this, I conduct an event study of the effect of having a brother on women's earnings percentile from age 18 through age 40 with age 18 as the base, controlling for individual fixed effects. Note, that sibling gender has a tight zero impact on earnings at age 18 (not reported). Once women enter the labor market,[22] we observe a negative effect of having a brother on the earnings percentile in the order of 0.5 percentile [Graph (a), Figure 2].[23] Such negative impact on earnings might be driven by differences in labor market participation rather than or in addition to occupational choice. This is, nevertheless, not the case, as Graphs (b) and (c) in Figure 2 illustrate that sibling gender does not affect women's cumulated work experience or cumulated length of unemployment. These findings of no effect on labor market participation (in terms of hours worked and unemployment) stress that labor market participation is not gendered in Denmark.

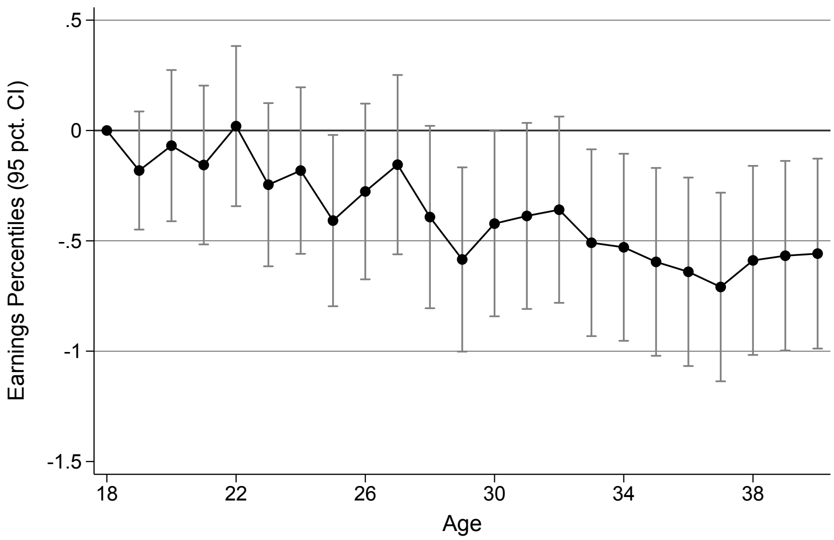
The finding of a negative consequence for earnings is not surprising, given the previous results of a lower participation in more male-dominated and STEM occupations. Similarly, Cools and Patacchini (2017) show that women in the U.S. with any brother earn less around age 30. Rao and Chatterjee (2017) do not find a significant effect of sibling gender composition on women's earnings among slightly older cohorts in the U.S., although their estimate of the effect of having a next younger male sibling indicates a negative impact. In contrast, studying a sample of female twins born in the first half of the last century, Peter et al. (2015) do not find an impact of having a co-twin brother on earnings. Moreover, both Cools and Patacchini (2017) and Rao and Chatterjee (2017) do not find significant effects of sibling gender composition on the

[22] Brenøe and Lundberg (2017) show, using Danish data, that almost everybody will have finished their education around age 30.

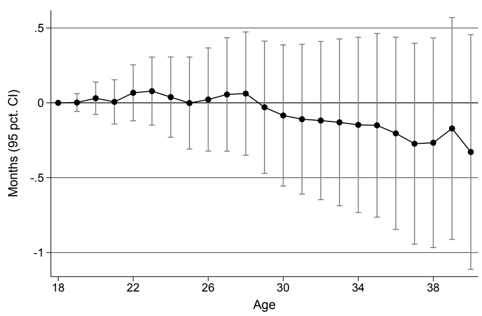
[23] Appendix Figure A3 demonstrates that the picture is similar when instead considering the earnings level and the natural logarithm of earnings.

Figure 2

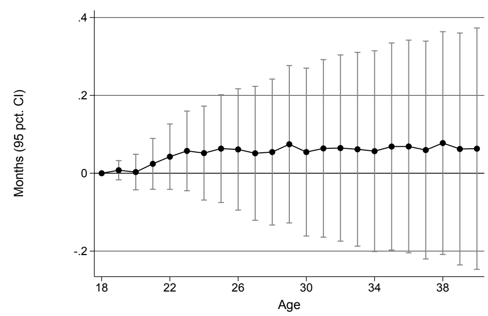
Effect of Sibling Gender on Labor Market Outcomes Age 18–40



(a) Earnings Percentile



(b) Work Experience



(c) Unemployment

Note: Main sample (first-born women born 1962–1975 with a younger biological sibling born within four years apart). The whiskers represent the 95 percent confidence interval. All graphs illustrate the estimates from an event study of the effect of having a second-born brother, where age 18 forms the base. All models absorb time-specific fixed effects and individual fixed effects. *Earnings Percentile* measures the labor earnings percentile by age and cohort. *Work Experience* measures the cumulated lifetime work experience in months. *Unemployment* measures the cumulated lifetime unemployment in months.

type of occupation. This might be due to some important empirical limitations (as the sign of their estimates broadly support my findings) because these studies rely on much smaller sample sizes, self-reported measures of occupation, and their methodological approach (i.e. the inclusion of all siblings in the measure of sibling gender composition and the inclusion of all birth orders in the sample).

4.3 Education and Family Formation

Another reason for the lower earnings could be due to differences in the accumulation of human capital. I do not find any evidence of an impact of sibling gender on educational attainment or school performance [Columns (2) and (3) in Panel A, Table 5].[24] Likewise, Cyron et al. (2017) does not find an effect of sibling gender on girls' cognitive or non-cognitive skills in first grade in the U.S.[25] Thus, sibling gender does not seem to affect differences in ability or (financial constraints in terms of) access to education. Consequently, these results demonstrate that sibling gender composition does not affect educational achievement or attainment, supporting an interpretation that changes in interests or identity the channels of the effects of sibling gender on occupational choice. In contrast, the only existing study with causal estimates of sibling gender on educational attainment finds that having a male co-twin increases women's length of education (Peter et al., 2015). However, their sample might not be comparable to the more general population of singletons and for later birth cohorts.

While sibling gender does not affect overall educational attainment, the effect of sibling gender on occupational choice is closely mirrored in field of education by age 30. Having a brother reduces the share of men in the highest completed field-by-level of education by 1.36 percent.[26] Similarly, women with a brother relative to those with a sister are respectively 7.6 and 11.3 percent less likely to ever enroll in and complete any field-specific STEM education. Appendix Table A5 further shows that the effect is present already in the type of first educational enrollment after compulsory education and that it is seen for STEM degree completion at different levels of education. Thus, having a brother pushes women out of traditionally male-dominated fields as early as age 16 and is both seen in field of education as well as occupation.

The magnitude of the effects are comparable to previous studies examining the impact of various aspects of the social environment in school on study choice (Bottia et al., 2015; Carrell et al., 2010; Schneeweis and Zweimüller, 2012; Fischer, 2017). Moreover, the results are broadly comparable to other studies examining correlations between

[24] Not reported, sibling gender does not affect the probability of having an observation on high school GPA or the probability of enrolling or completing different levels of education. Appendix Table A5 further shows that there is no effect on different types of ability, measured through grade 9 language and math written exam GPA. Appendix Figure A4 further illustrates the distributions of the three GPA measures by sibling gender composition. The differences by sibling gender are extremely small; thus, distributional effects do not seem to be important.

[25] Similarly, I do not find any effect of sibling gender on personality traits [Big Five, growth mindset, trust, hedonism] or mental health [Strength and Difficulties Questionnaire (SDQ)] (not reported), based on the DALSC sample introduced in Section 5.

[26] Despite large changes in society over time, the effect of sibling gender on the male share in the highest completed education by age 30 does not differ systematically by decade of birth when including cohorts born though 1986 (not reported). This is consistent with the finding by Haines et al. (2016) that gender-stereotypes have not changed over the last three decades in the U.S.

Table 5

Effect of Sibling Gender on Education and Family Formation

	(1)	(2)	(3)	(4)	(5)
Panel A: Education by age 30					
	Log(Male Share)	Length (months)	High School GPA	STEM Enrollment	STEM Completion
Second-Born	-1.36***	-0.12	-0.01	-0.63***	-0.57***
Brother	(0.53)	(0.15)	(0.01)	(0.17)	(0.13)
Observations	103,541	103,562	47,588	103,769	103,769
Panel B: Family Formation by age 41					
	Cohabit 18–41	Married 18–41	Has Any Children	# of Children	Age at First Birth
Second-Born	-0.80***	-0.12	-0.20	0.00	0.07**
Brother	(0.13)	(0.17)	(0.20)	(0.01)	(0.03)
Observations	103,769	103,769	103,769	103,769	91,953

Estimates in Columns (1), (4), and (5) in Panel A and Columns (1), (2), and (3) in Panel B are multiplied by 100 to express effects in percentage/log-points. Standard errors in parentheses; * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Main sample (first-born women born 1962–1975 with a second-born biological sibling born within four years apart). Each Panel-Column presents estimates from separate regressions. All models absorb fixed effects for birth municipality, year-by-month of birth, spacing in months to younger sibling, maternal age at birth, paternal age at birth, maternal level-by-field of education, and paternal level-by-field of education. The educational outcome models, except for high school GPA, further control for age at last observation in the education registry. *Log(Male Share)* measures the natural logarithm of the share of men in the highest completed education (narrow field-by-level) by age 30. *Length* measures the length of the highest completed education in months by age 30. *High School GPA* measures final GPA from the academic high school and is standardized by track and year of graduation for the total population with mean zero and standard deviation of one. *STEM Enrollment* indicates whether the woman has ever enrolled in a field-specific STEM education at age 16–27. *STEM Completion* indicates whether the woman has ever completed a field-specific STEM education by age 30. *Cohabit* measures the share of years age 18–41 during which the woman has cohabited with a partner without being married. *Married* measures the share of years age 18–41 during which the woman has been married. *Has Any Children* indicates whether the woman has at least one child by age 41. *# of Children* measures the number of children the woman has by age 41. *Age at First Childbirth* measures the age at the woman's first childbirth in years, conditional on having any children.

sibling gender composition and field of college major (Anelli and Peri, 2014; Oguzoglu and Ozbeklik, 2016). Appendix Table A10 displays the associations between gender of a first-born sibling and second-born women's gender identity, indicating similar but less robust correlations compared to the main results. These results are also closer to the ones in Anelli and Peri (2014) who do not find a significant association for women's enrollment in a high-earnings college major (although the magnitude of their estimate is relatively large). This stresses the importance of rigorously considering selection bias when the aim is to evaluate the causal effect of sibling gender.

In addition to differences in occupational and educational choice, one potential explanation for the negative effects on earnings might be differences in family formation. On one hand, due to the acquisition of more traditional gender norms, one might expect women with a brother to marry earlier, have children earlier, and have more children than women with a sister. However, such a conjecture implicitly requires that being married and having children is an important aspect of women's gender identity. This might very well not be the case in a modern setting in which women do not face a conclusive choice between having a family and a career (Bertrand et al., 2016; Goldin and Katz, 2002). The cohorts of women under study have, for instance, all had access to contraceptives, abortion, various family leave policies, and infant child-care options.[27] On the other hand, women with a younger sister might experience more competition in terms of being the first among the two who marries and has children, as men on average are older when they start their family formation. These two opposing forces might explain why I essentially do not find any effect of sibling gender on various aspects of family formation [Panel B in Table 5], consistent with the findings in Peter et al. (2015).

In terms of family formation, the results only suggest a small effect on cohabitation. Women with a brother cohabit 3.0 percent fewer years than those with a sister between age 18 and 41. This could, in fact, be due to more traditional gender norms, as more traditional women might want to wait longer before moving together with a partner before marriage.[28] Sibling gender has no effect on the probability of being married [Column (2)], age of first marriage, the probability of divorce, or age at first divorce (not reported). Thus, the only difference between women with a brother and those with a sister is that the former move together with a partner before marriage slightly later. This might explain the small positive (though negligible) effect on age at first childbirth. Overall, sibling gender has no effect on the fertility rate through age 41, i.e.

[27] Oral contraceptives (the pill) have been on the Danish market since 1966. All women have had free access to abortion since 1973 in Denmark.

[28] The vast majority of these cohorts cohabit before marriage and have children before marriage. Ninety-one percent of the women in the sample have cohabited at least one year before the year they get married and 53 percent get married in the year of their first childbirth or later.

close to complete realized fertility. Therefore, the effects of sibling gender on family formation are not a likely mediator of the effects on earnings. This, in turn, supports an interpretation of a causal positive effect of having a more male-dominated education and working in more male-dominated occupations on female earnings.

4.4 Persistent Effects to the Next Generation (of Girls)

Table 6

Effect of Sibling Gender on First-born Children's Grade 9 Performance

	Daughters		Sons	
	Language (1)	Math (2)	Language (3)	Math (4)
Second-Born Borther	2.37** (1.05)	0.19 (1.09)	0.36 (1.10)	0.33 (1.09)
Observations	29,047	29,036	29,262	29,262
Average	39.3	13.1	-6.0	23.8

All estimates are multiplied by 100 to express effects in percent of a standard deviation. Standard errors in parentheses; * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. First-born children to the main sample (first-born women born 1962–1975 with a second-born biological sibling born within four years apart) born 1986–1999. All models absorb fixed effects for the mother's birth municipality, year-by-month of birth, spacing in months to younger sibling, maternal age at birth, paternal age at birth, maternal level-by-field of education, and paternal level-by-field of education. The Grade 9 GPA measures come from the written exam at the end of grade 9 in respectively Danish and Math and are standardized by year of graduation for the total population with mean zero and standard deviation of one.

So far, I have documented that the childhood family environment affects the development of women's gender conformity. Having a brother influences the family environment to such a degree that women choose more female-dominated occupations and more gender-conforming partners. An intriguing question is whether this effect on gender identity is sufficiently strong to affect the next generation—and in particular, the next generation of girls. To investigate this, I examine the school performance of these women's first-born daughters and sons, separately. If having a more gender-stereotypical mother (and father) affects the next generation, we would expect daughters to perform better in languages and/or worse in math. For, boys the prediction is less clear, as the literature typically finds boys to be less sensitive to the social environment (Bottia et al., 2015; Carrell et al., 2010; Fischer, 2017). Remarkably, Table 6 shows that daughters whose mother's second-born sibling is male relative to female perform 2.37 percent of a standard deviation better in languages, while there is no effect on their math performance or for sons. Thus, daughters' difference in language and math ability is larger for those with a more gender-conforming mother. This increase in girls' absolute advantage in languages over math might, in turn, predict more traditional choices of field of education. Notably, I find evidence of very persistent long-run consequences of women's childhood family environment.

5 Gender-Specific Parenting as a Relevant Mechanism

5.1 Literature Background

The previous section documents that sibling gender does matter for women's acquisition of traditional gender norms and that the effects seem to be largest among women from more gender-stereotypical families. This subsection draws on the literature to identify relevant mechanisms behind these findings, while the subsequent subsection provides some empirical evidence. Overall, I consider changes in identity to be the main channel of the impacts on choice of occupation and parter, as the previous analysis does not suggest that differences in educational attainment, ability, labor force participation, family size, or resource constraints are important or driving mechanisms. Consistent with the same sex education literature (Booth et al., 2014; Schneeweis and Zweimüller, 2012), the overarching argument is that girls with a brother are more exposed to gender-stereotypical behavior in the family and are therefore more inclined to acquire traditional gender norms. In this context, gender-stereotypical behavior could become more salient through changes in the nature of either child-sibling or child-parent interactions, including parental investments.[29]

First, parents might interact differently with their children depending on the gender composition in terms of quantity, quality, and content of time spent together. Assuming that both parents spend at least some time with their children, a traditional household specialization model suggests that parents gender-specialize their investment in children when having mixed sex children if mothers are more productive in creating female human capital and fathers are more effective in creating male human capital (Becker, 1973). Parents might also derive more utility from spending time with a same compared to an opposite sex child due to the type of activities done with the child. In both cases, parents of mixed sex children would gender-specialize, to a greater extent, than parents of same sex children.

[29] Appendix A.4 provides a short overview of alternative mechanisms discussed in previous papers on sibling gender composition. These mechanisms cannot be the dominating explanations, as they are not compatible with the empirical findings.

McHale et al. (2003) suggest that because parents of mixed sex children have the opportunity to gender-differentiate their parenting, children with opposite gender siblings might have the strongest explicit gender-stereotypes. Endendijk et al. (2013) find some evidence that fathers with mixed sex children exhibit stronger gender-stereotypical attitudes than fathers with same sex children. Previous research has further documented that, overall, mothers talk more in general and more about interests and attitudes with daughters than sons (Maccoby, 1990; Leaper et al., 1998; Noller and Callan, 1990). Fathers, in contrast, talk more and spend more time with sons than daughters and have a greater emotional attachment to sons (Bonke and Esping-Andersen, 2009; Morgan et al., 1988; Noller and Callan, 1990). These different pieces of evidence thus suggest that parents of mixed sex children gender-specialize their parenting more and thereby expose their children more to gender-stereotypical behavior than parents of same sex children, which in turn might result in a stronger transmission of gender norms in families with mixed sex children.

Second, first-born girls might interact differently with their second-born sibling depending on the siblings' gender. In particular, having a brother might make girls more aware of "appropriate" female behavior and thereby induce them to develop more gender-stereotypical attitudes. For instance, Booth and Nolen (2012) show that girls attending same sex schools are no more risk averse than boys, while girls attending mixed sex schools are significantly more risk averse. Women are generally less competitive than men and this gender difference in competitiveness seems to be larger in mixed sex relative to same sex environments (Bertrand, 2011; Niederle and Vesterlund, 2011). Traditionally male-dominated (STEM) fields are further considered more competitive (Buser et al., 2014). Therefore, having a brother instead of a sister might change women's degree of competitiveness and thereby their preferences for working in competitive environments. Having a brother might thereby induce women to develop more gender-stereotypical attitudes due to a greater awareness of gender through sibling interactions. This, in turn, could be reinforced by parents' increased gender-specialization. In particular, previous studies have documented that women with brothers behave more family-centered and express more traditional attitudes towards gender roles (Cools and Patacchini, 2017; Rao and Chatterjee, 2017).

Thus, a particularly important mechanism for the observed effect of sibling gender on women's formation of gender identity—that I am able to test for empirically—is differences in child-parent interactions and, in particular, increased gender-specialization in families with mixed sex children. In the remainder of this section, I explore this mechanism by investigating the impact of sibling gender composition on parental time investment. More precisely, in the daily child-parent interactions, we might observe that parents of mixed sex children invest more quality time in their same sex child. This could explain the heterogeneity in the effect of sibling gender documented in Table 4. Furthermore, in the case of parental divorce, we might expect that children from mixed sex child-families would be more likely to live with their same sex parent compared to same sex children due to a larger degree of gender-specialized parenting. Consequently, common for these predictions is that a parent of mixed sex children influences his or her same sex child more than a parent of same sex children.

5.2 Empirical Evidence on Gender-Specific Parenting

To investigate whether sibling gender composition affects child-parent interactions, I draw on the Danish Longitudinal Survey of Children (DALSC).[30] The survey consists of five waves of children born in 1995 and is unique due to its very detailed information on parental time use and family socio-economic characteristics. For this analysis, I select first-born girls who have a second-born sibling born within four calendar years apart.[31] At age 7 and 11, both parents report how often they do different types of activities together with their first-born daughter. I construct an index on parental time investment, using principal component analysis, and standardize it with mean zero and standard deviation of one [Appendix Table A6]. I define quality time as playing with the child, helping with homework, doing out-of-school activities, reading/singing, and going on an excursion.

Columns (1) through (4) in Table 7 provide the results on parental time investment by each parent for the two ages, separately. Mothers of a first-born daughter and a second-born son invest *more* time in their first-born daughter at both ages compared to mothers with two daughters. The increase is in the magnitude of 14–17 percent of a standard deviation. In contrast, fathers invest 20–23 percent of a standard deviation *less* time in their first-born daughter when having mixed sex children. This reduction in total paternal time investment is driven by decreased time spent helping with homework and reading for the daughter [Appendix Table A7]. This finding indicates that girls with a younger brother receive less qualified help with homework in traditionally male-dominated subjects, which might prevent them from growing interests in these fields. This effect on father-daughter interactions furthermore translates into a substantially worse relationship between fathers and their first-born daughters when the second-born child is male relative to female [Appendix Table A8]. Overall, girls receive the same amount of time investment regardless of their younger sibling's gender. These results clearly show that first-born girls with a second-born brother experience more gendered

[30] The study was designed by researchers from SFI, the Danish National Centre for Social Research, in collaboration with other research institutions. The survey consists of 6,011 randomly sampled children born between September and October, 1995 to a mother with Danish citizenship and consists of five waves (1996, 1999, 2003, 2007, and 2011).

[31] I only observe the year of birth of siblings and do therefore not have more precise information on spacing.

Table 7

Effect of Sibling Gender on Parental Time Investment in First-Born Daughters and Family Structure

	Parental Time Investment (Born 1995)				Family Structure (Born 1962–75)	
	Mother		Father		Lives w Both Parents	Lives w Mother & Sib w Father
	Age 7	Age 11	Age 7	Age 11		
	(1)	(2)	(3)	(4)	(5)	(6)
Second-Born Brother	0.14*	0.17**	-0.20**	-0.23**	0.11	5.30***
	(0.08)	(0.09)	(0.10)	(0.10)	(0.23)	(0.38)
Observations	594	562	421	415	102,137	19,196
Average	-0.00	0.00	-0.00	0.00	81.1	7.3
DALSC Sample	✓	✓	✓	✓		
Main Sample						
All					✓	
Divorced						✓

Standard errors in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. DALSC sample: Columns (1) through (4). Main sample: Columns (5) and (6). Each Column represents the results from separate regressions. All models using the DALSC sample control for (quadratic) mother and father's age and fixed effects for spacing to the younger sibling in years, parental marital status in 1996, parents having been together for at least 5 years in 1996, region of birth, maternal level of education, paternal level of education, and family income level in 1995. Both models using the main sample absorb fixed effects for birth municipality, year-by-month of birth, spacing in months to younger sibling, maternal age at birth, paternal age at birth, maternal level-by-field of education, paternal level-by-field of education, and age at observation of family structure. *Parental time investment* is constructed, using principal component analysis based on reports on how often each parent does certain quality time activities (playing, doing homework, doing out-of-school activities, reading/singing, going on an excursion) together with the child at a weekly basis and is standardized with mean zero and standard deviation of one. *Main Sample All* includes everybody who lives with at least one biological parent, while *Main Sample Divorced* excludes those living with both biological parents. *Lives w Both Parents* indicates that the first-born daughter lives with both biological parents at age 17. *Lives w Mother & Sib w Father* indicates that the first-born daughter lives with her mother and the second-born child lives with the father at age 17.

parenting relative to those with a younger sister.[32]

Ideally, I would have had similarly detailed data on parental inputs for the main sample. Such information is, however, not observed in the administrative registries. Instead, I do observe all children's family structure at age 17.[33] Sibling gender composition does not alter the probability of living with both biological parents [Column (5) in Table 7]. In the case of parental divorce or separation (henceforth *divorce*), the living arrangement between parents and children might additionally help shed light on child-parent interactions in terms of splitting parents' time in the main sample. If parents of mixed sex children gender-specialize more than parents of same sex children, we would expect that divorced families with mixed sex children would be more likely than families with same sex children to have a living arrangement in which the first-born daughter lives with her mother and the second-born child lives with the father.

Conditional on living in a divorced family, the results show a pattern consistent with the prediction [Column (6)]. First-born daughters with a second-born brother are 5.30 percentage points (115 percent) more likely to live with their mother while their younger sibling lives with the father. These results consequently show a strong effect on the living arrangement among non-traditional families, thereby lending support to the previous findings (based on the much smaller DALSC sample) on more gender-specific parenting and time investment in families with mixed sex children. In conclusion, these findings support the hypothesis that parents of mixed sex children gender-specialize their parenting more than parents of same sex children, thereby strengthening the transmission of traditional gender-specific interests.[34]

6 First-Born Men and their Second-Born Sisters

The main analysis investigates the effect of sibling gender on the origins of women's gender identity. This section briefly presents a corresponding analysis for men. How-

[32]For first-born boys, the overall picture is similar (not reported). Note, I cannot distinguish whether this increase in gender-specialization is driven by changes in demand (children) or supply (parents). Having a brother might cause the daughter to demand more maternal and less paternal time. The results, however, clearly show that parents respond to sibling gender, which is the relevant margin, as any policy aiming at reducing the transmission of gender norms would most likely need to address parents and not children as young as 7 years.

[33]I observe the family structure on January 1^{st} each year and use the observation for the year the person turns 18 years or the last year in which the child lives with at least one biological parent.

[34]Not reported, considering heterogeneity by living in a traditional family for occupational choice shows that the effect is largest for women from divorced families. This is consistent with increased gender-specialization in these families. However, there is no significant heterogeneity by family structure for working in STEM occupations or choice of partner.

ever, I do not consider men's choice of partner or the school performance of their first-born children, because I find that sibling gender affects men's family formation both in terms of having a partner and having any children [Panel B in Appendix Table A12]; put differently, considering those outcomes might create selection issues and potentially bias the estimates. I construct the sample of men with identical selection criteria as for the main sample of women and conduct an identical analysis with the same variable definitions and controls.

Overall, the results for first-born men suggest that having a second-born sister relative to a second-born brother enhances men's gender identity [Appendix Table A11]. Men with a sister have a slightly higher (borderline significant) share of men in their occupation and are 0.51 percentage points (3.7 percent) more likely to work within STEM.[35] Importantly, however, having a sister also decreases the probability of working in managerial occupations by 0.44 percentage points (6.6 percent).[36] This decrease in the likelihood of working in (high-paid) managerial positions may help explain why men with a sister experience lower labor earnings than men with a brother [Appendix Figure A5]. At the same time, men with a sister cumulate less work experience at the end of their 30s relative to those with a brother, while there is no effect on lifetime unemployment by age 40. Thus, men with a sister appear somehow less successful in the labor market.

Similar to my findings, previous studies find negative effects of having sisters relative to brothers on men's earnings in Sweden and the U.S. (Peter et al., 2015; Rao and Chatterjee, 2017). Rao and Chatterjee (2017) show that in the U.S. brothers help each other more in job search than mixed sex siblings, which could help explain the negative effect on earnings and be a mechanism counteracting our ability to observe men's gender identity through occupational choice. Moreover, Peter et al. (2015) discuss competition between brothers as an important channel of the positive effect of having a brother on earnings. Brothers might compete with each other to a much greater extent than mixed sex siblings, both because men are more competitive than women and because having a same sex sibling might change the reference point of competition (Butcher and Case, 1994; Conley, 2000). Joensen and Nielsen (2017) show that especially brother pairs influence each other in terms of educational choice. Panel A in Appendix Table A12 shows that having a sister increases men's probability of ever enrolling in any field-specific STEM (traditionally heavily male-dominated) programs, supporting a change in their gender identity. However, the effect does not persist into actual degree completion, which again may suggest that having a sister decreases com-

[35] The results are comparable when considering a binary indicator for having ever worked in STEM from age 31 through 40 (not reported).

[36] Not reported, I find a tight zero effect of sibling gender on women's probability of working in managerial occupations (the estimated effect is 0.04 percentage points ($se = 0.07$)).

petitive behavior, making them strive—and in the end—achieve less. Besides the effect on STEM enrollment, sibling gender does not impact men's educational attainment or achievement.

Like Peter et al. (2015), I also find that having a sister affects men's family formation negatively. Men with a sister cohabit and are married fewer years from age 18 through 41. Furthermore, having a sister decreases men's probability of having any children and their number of children. These findings could reflect less competitive behavior among men with a sister relative to those with a brother not only in the labor market but also in the marriage market. Thus, despite finding indications of similar effects of having an opposite sex sibling on men's development of gender norms as for women, competition might play a similarly or more important role for how men fare in the labor and marriage markets.

7 Conclusion

This study documents that the childhood family environment has a long-run impact on women's gender identity with persistent effects to the next generation of girls. The results show that having a second-born brother relative to a sister increases first-born women's gender conformity, both in terms of their choice of occupation and partner. I further show that having a brother negatively affects labor earnings. This is most likely driven by the effect on occupational choice, as sibling gender does not affect educational attainment, labor market participation, or family formation. I provide compelling evidence that changes in child-parent interactions—and, in particular, increased gender-specialized parenting in families with mixed sex children—play an important role for the changes in gender identity. This suggests that the transmission of traditional gender norms is stronger in families with mixed sex children. Finally, I show that the increased gender conformity among women with a brother persists into the next generation of girls, as indicated by an increase in daughters' comparative advantage in language over math performance in school. Consequently, I find evidence of very persistent long-run consequences of women's childhood family environment.

To eliminate gender inequality caused by gender-conforming behavior, my findings imply that policy makers need to focus on the formation of gender identity among girls in the childhood family environment. I show that having a brother affects girls' study choice in a more gender-stereotypical direction already at the end of compulsory schooling. This stresses that girls' development of gender identity by adolescence has important consequences for their later-life educational and labor market outcomes. As my mechanism analysis suggests, the family—representing a central aspect of the social environment during childhood—influences the formation of women's gender identity.

Therefore, if society wants to give boys and girls the same opportunities at the time they enter the labor market in adulthood, policy makers would need to focus on how to counteract gender-stereotypical human capital investments. Specifically, interventions would need to counteract the transmission of gender norms across generations and thereby the development of gender-stereotypical behaviors, attitudes, and preferences.

References

Akerlof, G. A. and R. E. Kranton (2000). Economics and identity. *Quarterly journal of Economics*, 715–753.

Alan, S., S. Ertac, I. Mumcu, et al. (2017). Gender stereotypes in the classroom and effects on achievement.

Altonji, J. G., P. Arcidiacono, and A. Maurel (2015). The analysis of field choice in college and graduate school: Determinants and wage effects.

Amin, V. (2009). Sibling sex composition and educational outcomes: A review of theory and evidence for the UK. *Labour 23*(1), 67–96.

Andersson, G., K. Hank, M. Rønsen, and A. Vikat (2006). Gendering family composition: Sex preferences for children and childbearing behavior in the nordic countries. *Demography 43*(2), 255–267.

Anelli, M. and G. Peri (2014). Gender of siblings and choice of college major. *CESifo Economic Studies 61*(1), 53–71.

Anelli, M. and G. Peri (2016). The effects of high school peers' gender on college major, college performance and income.

Angrist, J., V. Lavy, and A. Schlosser (2010). Multiple experiments for the causal link between the quantity and quality of children. *Journal of Labor Economics 28*(4), 773–824.

Angrist, J. D. and W. N. Evans (1998). Children and their parents' labor supply: Evidence from exogenous variation in family size. *The American Economic Review 88*(3), 450.

Bauer, T. and I. Gang (2001). Sibling rivalry in educational attainment: The german case. *Labour 15*(2), 237–255.

Becker, G. S. (1973). A theory of marriage: Part i. *Journal of Political economy 81*(4), 813–846.

Bertrand, M. (2011). New perspectives on gender. *Handbook of labor economics 4*, 1543–1590.

Bertrand, M., P. Cortés, C. Olivetti, and J. Pan (2016). Social norms, labor market opportunities, and the marriage gap for skilled women.

Black, S. E., P. J. Devereux, and K. G. Salvanes (2005). The more the merrier? the effect of family size and birth order on children's education. *The Quarterly Journal of Economics 120*(2), 669–700.

Black, S. E., E. Grönqvist, and B. Öckert (2017). Born to lead? the effect of birth order on non-cognitive abilities. *Review of Economics and Statistics* (0).

Blau, F. D. and L. M. Kahn (2016). The gender wage gap: Extent, trends, and explanations.

Bonke, J. and G. Esping-Andersen (2009). Parental investments in children: how educational homogamy and bargaining affect time allocation. *European Sociological Review 10*(20), 1–13.

Booth, A., L. Cardona-Sosa, and P. Nolen (2014). Gender differences in risk aversion: Do single-sex environments affect their development? *Journal of economic behavior & organization 99*, 126–154.

Booth, A. L., L. Cardona, and P. J. Nolen (2013). Do single-sex classes affect exam scores? an experiment in a coeducational university.

Booth, A. L. and P. Nolen (2012). Gender differences in risk behaviour: does nurture matter? *The Economic Journal 122*(558).

Bottia, M. C., E. Stearns, R. A. Mickelson, S. Moller, and L. Valentino (2015). Growing the roots of stem majors: Female math and science high school faculty and the participation of students in stem. *Economics of Education Review 45*, 14–27.

Brenøe, A. A. and T. Epper (2018). The Intergenerational Transmission of Time Preferences Persists across Four Decades.

Brenøe, A. A. and S. Lundberg (2017). Gender gaps in the effects of childhood family environment: Do they persist into adulthood? *European Economic Review*.

Brenøe, A. A. and R. Molitor (2018). Birth Order and Health of Newborns. *Journal of Population Economics 31*(2), 363–395.

Brenøe, A. A. and U. Zölitz (2018). Exposure to More Female Peers Widens the Gender Gap in STEM Participation.

Brim, O. G. (1958). Family structure and sex role learning by children: A further analysis of helen koch's data. *Sociometry 21*(1), 1–16.

Buser, T., M. Niederle, and H. Oosterbeek (2014). Gender, competitiveness and career choices. *The Quarterly Journal of Economics 129*(3), 1409–1447.

Butcher, K. F. and A. Case (1994). The effect of sibling sex composition on women's education and earnings. *The Quarterly Journal of Economics 109*(3), 531–563.

Carlana, M. (2017). Stereotypes and self-stereotypes: Evidence from teachers gender bias.

Carrell, S. E., M. E. Page, and J. E. West (2010). Sex and science: How professor gender perpetuates the gender gap. *The Quarterly Journal of Economics 125*(3), 1101–1144.

Cheng, A., K. Kopotic, and G. Zamarro (2017). Can parents growth mindset and role modelling address stem gender gaps? *EDRE Working Paper*.

Conley, D. (2000). Sibship sex composition: Effects on educational attainment. *Social Science Research 29*(3), 441–457.

Cools, A. and E. Patacchini (2017). Sibling Gender Composition and Women's Wages. *IZA Working Paper* (11001).

Cronqvist, H., A. Previtero, S. Siegel, and R. E. White (2015). The fetal origins hypothesis in finance: Prenatal environment, the gender gap, and investor behavior. *The Review of Financial Studies 29*(3), 739–786.

Cyron, L., G. Schwerdt, and M. Viarengo (2017). The effect of opposite sex siblings on cognitive and noncognitive skills in early childhood. *Applied Economics Letters*, 1–5.

Endendijk, J. J., M. G. Groeneveld, S. R. van Berkel, E. T. Hallers-Haalboom, J. Mesman, and M. J. Bakermans-Kranenburg (2013). Gender stereotypes in the family context: Mothers, fathers, and siblings. *Sex Roles 68*(9-10), 577–590.

Farré, L. and F. Vella (2007). The intergenerational transmission of gender role atti-

tudes and its implications for female labor force participation.
Fernández, R., A. Fogli, and C. Olivetti (2004). Mothers and sons: Preference formation and female labor force dynamics. *The Quarterly Journal of Economics 119*(4), 1249–1299.
Fischer, S. (2017). The downside of good peers: How classroom composition differentially affects men's and women's stem persistence. *Labour Economics*.
Gielen, A. C., J. Holmes, and C. Myers (2016). Prenatal testosterone and the earnings of men and women. *Journal of Human Resources 51*(1), 30–61.
Goldin, C. (2014). A grand gender convergence: Its last chapter. *American Economic Review 104*(4), 1091–1119.
Goldin, C. and L. F. Katz (2002). The power of the pill: Oral contraceptives and womens career and marriage decisions. *Journal of political Economy 110*(4), 730–770.
Haines, E. L., K. Deaux, and N. Lofaro (2016). The times they are a-changing or are they not? a comparison of gender stereotypes, 1983–2014. *Psychology of Women Quarterly*, 0361684316634081.
Hauser, R. M. and H.-H. D. Kuo (1998). Does the gender composition of sibships affect women's educational attainment? *Journal of Human Resources*, 644–657.
Humlum, M. K., A. Nandrup, and N. Smith (2017). Closing or reproducing the gender gap? parental transmission, social norms and education choice.
Joensen, J. S. and H. S. Nielsen (2017). Spillovers in education choice. *Journal of Public Economics*.
Johnston, D. W., S. Schurer, and M. A. Shields (2013). Maternal gender role attitudes, human capital investment, and labour supply of sons and daughters. *Oxford Economic Papers 66*(3), 631–659.
Kaestner, R. (1997). Are brothers really better? sibling sex composition and educational achievement revisited. *Journal of Human Resources 32*(2).
Kahn, S. and D. Ginther (2017). Women and stem. Technical report, National Bureau of Economic Research.
Kirkebøen, L., E. Leuven, and M. Mogstad (2016). Field of study, earnings, and self-selection. *The Quarterly Journal of Economics*, qjw019.
Kleven, H., C. Landais, and J. E. Søgaard (2018). Children and gender inequality: Evidence from denmark.
Koch, H. L. (1955). Some personality correlates of sex, sibling position, and sex of sibling among five-and six-year-old children. *Genetic Psychology Monographs*.
Lavy, V. and R. Megalokonomou (2017). Persistency in teachers grading biases and effect on longer term outcomes: university admission exams and choice of field of study. *Unpublished Manuscript*.
Lavy, V. and E. Sand (2015). On the origins of gender human capital gaps: Short and long term consequences of teachers stereotypical biases.
Leaper, C., K. J. Anderson, and P. Sanders (1998). Moderators of gender effects on parents' talk to their children: a meta-analysis. *American psychologist 34*(1), 3–27.
Lehmann, J.-Y. K., A. Nuevo-Chiquero, and M. Vidal-Fernandez (2016). The early origins of birth order differences in childrens outcomes and parental behavior. *Journal of Human Resources*, 0816–8177.
Maccoby, E. E. (1990). Gender and relationships: A developmental account. *American psychologist 45*(4), 513–520.

McHale, S. M., A. C. Crouter, and S. D. Whiteman (2003). The family contexts of gender development in childhood and adolescence. *Social development 12*(1), 125–148.

McHale, S. M., K. A. Updegraff, and S. D. Whiteman (2013). Sibling relationships. In *Handbook of marriage and the family*, pp. 329–351. Springer.

Morgan, S. P., D. N. Lye, and G. A. Condran (1988). Sons, daughters, and the risk of marital disruption. *American journal of sociology 94*(1), 110–129.

Niederle, M. and L. Vesterlund (2011). Gender and competition. *Annu. Rev. Econ. 3*(1), 601–630.

Noller, P. and V. J. Callan (1990). Adolescents' perceptions of the nature of their communication with parents. *Journal of Youth and Adolescence 19*(4), 349–362.

OECD (2016). *Education at a Glance 2016: OECD Indicators*.

OECD (2017). Oecd labour force statistics 2017.

Oguzoglu, U. and S. Ozbeklik (2016). Like father, like daughter (unless there is a son): Sibling sex composition and womens stem major choice in college.

Olivetti, C., E. Patacchini, Y. Zenou, et al. (2016). Mothers, peers and gender identity. *Boston College, Department of Economics, Working Paper 904*.

Peri, G., K. Shih, and C. Sparber (2015). Stem workers, h-1b visas, and productivity in us cities. *Journal of Labor Economics 33*(S1), S225–S255.

Peter, N., P. Lundborg, and D. H. Webbink (2015). The effect of sibling's gender on earnings, education and family formation.

Powell, B. and L. C. Steelman (1989). The liability of having brothers: Paying for college and the sex composition of the family. *Sociology of Education*, 134–147.

Rao, N. and T. Chatterjee (2017). Sibling gender and wage differences. *Applied Economics*, 1–21.

Schneeweis, N. and M. Zweimüller (2012). Girls, girls, girls: Gender composition and female school choice. *Economics of Education review 31*(4), 482–500.

Zölitz, U. and J. Feld (2017). The effect of peer gender on major choice. *Mimeo*.

A Appendix

A.1 The Selection Bias Problem

To show the selection bias problem more formally, I here follow Peter et al. (2015). Assume a latent outcome $Y_i^* = \alpha + \beta G_i^{old} + X_i' \gamma + \epsilon_i$, where G_i^{old} is the gender of the older sibling and X_i is a vector of observable exogenous characteristics. ϵ_i contains other relevant unobservable variables, such as parental gender preferences denoted by P_i, and $E[\epsilon_i] = 0$. The bias arises because of the latent nature of Y_i^*, as we only observe the outcome if child i is born. In other words, $Y_i = Y_i^*$ if the child is born ($S_i = 1$) and Y_i is missing if the child is not born ($S_i = 0$). The selection depends both on parental preferences and the older child's gender, $S_i = f(P_i, G_i^{old})$. We can only estimate the effect for the sample of children who are born which gives the expected value of Y_i:

$$E[Y_i|S_i = 1, G_i^{old}, X_i] = \alpha + \beta G_i^{old} + \gamma X_i + E[\epsilon_i|S_i = 1, G_i^{old}, X_i] \tag{2}$$

$$= \alpha + \beta G_i^{old} + \gamma X_i + E[\epsilon_i|f(P_i, G_i^{old}) = 1, G_i^{old}, X_i].$$

As long as selection depends on the first child's gender and parental preferences affect the way in which parents raise their children $E[\epsilon_i|f(P_i, G_i^{old}) = 1, G_i^{old} = 1, X_i] \neq E[\epsilon_i|f(P_i, G_i^{old}) = 1, G_i^{old} = 0, X_i]$. This implies that the estimate of the older sibling's gender is biased.

A selection problem could also arise in the absence of parental gender preferences. Assume that first-born children have n normally-distributed traits, such as how easy the child is to take care of and how well it behaves. Suppose parents only want a second child if their first child has a value of each trait above a certain threshold. The threshold for or the distribution of each trait could be gender-specific. In both cases, parents who progress to the next parity would, on average, have different types of first-born children depending on the child's gender. For instance, if boys and girls have the same distribution of how well they behave but parents require girls to behave better than boys to have a second child, second-born children would, on average, have a better behaving older sibling if they have a sister compared to a brother. In this example, the estimated effect of the older sibling's gender on the younger child's outcomes might thus be due to the older sibling's behavior rather than due to his or her gender.

A.2 Family Size

Parents in developed countries are more likely to have a third child if their first two children are of same compared to mixed gender (Angrist and Evans, 1998; Angrist et al., 2010; Black et al., 2005). Appendix Table A1 shows that this is also the case in the main sample of the analysis. First-born women with a second-born brother are 13.2 percent less likely to have at least two siblings relative to those with a sister. The rest of this appendix examines whether family size has an independent effect on gender identity and studies rigorously the robustness of the main results to family size.

Table A1
Effect of Sibling Gender on Parental Realized Fertility

	# of Siblings (1)	≥ 2 Siblings (2)	≥ 3 Siblings (3)
Second-Born	-0.07***	-5.26***	-1.33***
Brother	(0.01)	(0.28)	(0.16)
Observations		103,769	
Average	1.6	37.2	7.7

Estimates for the outcomes ≥ 2 *Siblings* and ≥ 3 *Siblings* are multiplied by 100 to express effects in percentage points. Standard errors in parentheses; * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Main sample (first-born women born 1962–1975 with a second-born biological sibling born within four years apart). Each Column presents estimates from separate regressions. All models absorb fixed effects for birth municipality, year-by-month of birth, spacing in months to younger sibling, maternal age at birth, paternal age at birth, maternal level-by-field of education, and paternal level-by-field of education. *# of Siblings* measures the total number of siblings the individual has, including full and half siblings. $\geq 2(3)$ *Siblings* takes the value one if the person has at least two (three) full siblings and zero otherwise.

A.2.1 Does Family Size affect Gender Identity?

Black et al. (2005) use twins as an instrument for family size to show that family size does not affect educational attainment, using Norwegian registry data; Angrist et al. (2010) find the same for Israel. However, they only consider length of schooling and not gender identity. In this supplementary analysis, I show, consistent with their findings, employing a similar strategy in the Danish context, that family size does not affect educational attainment or the measures of gender identity used in the main analysis.

I use a sample with similar sample restrictions as for the main sample (see Subsection 3.1) with the exception that I include first-born singleton children who have younger twin siblings born at the second parity.[37] The instrument for family size is having twins at the second parity. Column (1) in Appendix Table A2 shows that the instrument is strong and relevant; see Angrist et al. (2010) and Black et al. (2005) for a discussion of the validity of the instrument.

Columns (2) through (6) show the second stage results. Similar to the findings for Norway and Israel, family size does not affect the length of highest completed

[37] I include all multiple births; twins, however, represent the vast majority of all multiple births.

Table A2

The Effect of Family Size on Gender Conformity using Twins as Instrument

	First Stage	Second Stage				
		Choice of Occ and Partner			Education	
	# of Siblings	Log(Male Share in own Occ)	Works in STEM	Log(Female Share in Partner's Occ)	Log(Male Share in Edu)	Length (months)
	(1)	(2)	(3)	(4)	(5)	(6)
Twins at 2^{nd} parity	0.71***					
	(0.02)					
# of Siblings		3.82	0.70	-1.85	-1.33	0.27
		(3.35)	(0.82)	(4.84)	(3.78)	(1.06)
F-statistic of IV	1020.11					
Prob>F	< 0.001					
Observations	104,780	104,780	104,780	95,977	104,552	104,573
Effect×-0.07		-0.27	-0.05	0.13	0.09	-0.02

All second stage estimates (except Length of Education) are multiplied by 100 to express effects in percentage/log-points. Standard errors in parentheses; * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Main sample including twin siblings born at second parity (first-born women born 1962–1975 with a second-born biological sibling born within four years apart). Each Column presents estimates from separate regressions. All models absorb fixed effects for birth municipality, year-by-month of birth, spacing in months to younger sibling, maternal age at birth, paternal age at birth, maternal level-by-field of education, and paternal level-by-field of education. *# of Siblings* measures the total number of siblings the individual has, including full and half siblings. Columns (2) and (3) also include dummies for the number of years observed in the income registry from age 31–40 and the number of years observed with a valid occupation code from age 31–40. Column (4) also includes dummies for the partner's number of occupational observations and age at first and last observation. The occupational outcomes of the first-born women are measured as mean from age 31–40. The occupational outcome of the partner is measured mainly at ages 31–45 for the partner with whom the woman lived most years from age 31–41. The effects are multiplied by -0.07 (*Effect×-0.07*), as it is the magnitude of the effect of having a brother on the number of siblings.

education by age 30. Neither does it significantly impact the womans' occupational choice, her choice of partner, or her type of education. The last row in the table scales the estimates by -0.07 (i.e. the effect of having a second-born brother on the total number of siblings). This statistic (*Effect*× *0.07*) illustrates that if family size (despite not having any statistically significant effect on the outcomes) would mediate some of the effect of sibling gender, any potential bias would be tiny.

A.2.2 Robustness to Family Size

Table A3

Splitting Sample by Family Size

	Log(Male Share in Own Occ)		Share of Years in STEM Occupation		Log(Female Share in Partner's Occ)	
	(1)	(2)	(3)	(4)	(5)	(6)
Second-Born	-1.09*	-1.11	-0.49***	-0.37**	-2.26**	-1.84
Brother	(0.62)	(0.82)	(0.16)	(0.18)	(0.88)	(1.12)
Observations	58313	36010	58313	36010	53148	33331
Average	788.4	784.9	5.5	4.4	299.3	290.7
# of Siblings	1	≥ 2	1	≥ 2	1	≥ 2

All estimates are multiplied by 100 to express effects in percentage/log-points. Standard errors in parentheses; * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Main sample with only full siblings (first-born women born 1962–1975 with a second-born biological sibling born within four years apart). Each Panel-Column presents estimates from separate regressions. All models absorb fixed effects for birth municipality, year-by-month of birth, spacing in months to younger sibling, maternal age at birth, paternal age at birth, maternal level-by-field of education, and paternal level-by-field of education. Columns (1) through (4) also include dummies for the number of years observed in the income registry from age 31–40 and the number of years observed with a valid occupation code from age 31–40. Columns (5) and (6) also includes dummies for the partner's number of occupational observations and age at first and last observation. The occupational outcomes of the first-born women are measured as mean from age 31–40. The occupational outcome of the partner is measured mainly at ages 31–45 for the partner with whom the woman lived most years from age 31–41. *1 Sibling*-models restrict the sample to those who only have one full sibling and no half-siblings. ≥ 2 *Siblings*-models restrict the sample to those who have at least two full siblings and no half-siblings.

As shown in Appendix Tables A1 and A3, sibling gender composition affects family size but family size does not affect gender identity. To further test the robustness of the main results to family size (in addition to flexibly control for family size as done in Column (4) in Table 3 in the main text), this subsection employs two alternative strategies: 1) to divide the sample by family size and 2) to study the effect of having a co-twin brother. Although family size is endogenous to sibling gender composition, strategy (1) is useful to the degree that it informs about the sensitivity of the results. These robustness analyses, together with the evidence of no differential effect by sibling gender on educational attainment or labor market participation [Table 5 and Figure 2] and the absence of an effect of family size on gender conformity, provide convincing evidence that family size does not confound the effects of sibling gender composition.

Table A4
Effect of Having A Co-Twin Brother on Gender Conformity in Education

	Next Birth	Log(Male Share in Edu)	STEM Enrollment	STEM Completion
	(1)	(2)	(3)	(4)
Co-Twin Brother	-1.27*	-4.23**	-1.64***	-1.50***
	(0.73)	(2.04)	(0.58)	(0.43)
Observations	9,380	9,357	9,380	9,380
Average	28.9	331.7	7.3	4.2

All estimates are multiplied by 100 to express effects in percentage/log points. Standard errors in parentheses, clustered at the mother level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Each Column presents estimates from separate regressions. The sample consists of twins born 1962–86. All models absorb fixed effects for birth county, year of birth, mother's level and field of education, father's level and field of education, parity, and age at last educational observation. The models further control for (cubed) mother's age at birth and (cubed) father's age at birth. *Next Birth* indicates if the parents get a subsequent child. *Log(Male Share in Edu)* measures the natural logarithm of the male share in the highest completed education (narrow field-by-level) by age 30. *STEM Enrollment* indicates whether the woman has ever enrolled in a field-specific STEM education at age 16–27. *STEM Completion* indicates whether the woman has ever completed a field-specific STEM education by age 30.

The first strategy is to split the sample by family size. For this, I restrict the sample to individuals who only have biological siblings, i.e. none of their parents have children with another person than the parent; though the results are similar when including those with half-siblings. Given family size is endogenous, this robustness check comes with a selection problem. If those parents of same sex children (born at the first two parities) who have a third child are more gender-stereotypical and to a greater extent influence their children's outcomes in such direction than those who do not have a third child, we would expect the effect of having a second-born brother to be larger in magnitude among first-born children from two-child families than for the entire sample. Similarly we would expect the effect of sibling gender to be smaller among children from families with at least three children. This is exactly what the results show in Table A3.

Finally, to circumvent potential confounding effects from family size, I examine the effect of having a co-twin brother as an alternative empirical strategy. This approach is similar to the one in Cronqvist et al. (2015) and Peter et al. (2015), except that I do not have information on zygocity. To increase power, I include birth cohorts 1962–1986 and consider the gender conformity in educational outcomes. The key empirical feature of the sample of twins is that twin gender composition only has a very limited impact on family size [Appendix Table A4, Column (1)]. Overall, the effects of having a co-twin brother on educational choice are similar to the main results. The magnitude of the effects are, however, much larger. This may be due to the much greater intensity of the exposure to a co-twin compared to a younger sibling.

A.3 Educational System and Field of Study

Throughout, I follow the International Standard Classification of Education (ISCED) for the definition of all educational measures. I include observations through age 27 for all enrollment measures and through age 30 for all completion measures to give people time to complete the education in which they enroll. I define the *male share in education* as the share of men who had their highest completed education at age 30 within the same narrow field and level of education for cohorts born 1–5 years before the individual. The academic high school grade point average (GPA) is standardized with mean zero and standard deviation of one at the year of graduation and high school track level for the total population; note, however, that it is only observed for those completing the academic high school.

In the final year of 9^{th} grade, at age 16, students decide whether to apply for secondary education or to enter the labor market.[38] Secondary education (ISCED level 3) consists of two types: academic high school and vocational training. The academic high school is generic (i.e. not field-specific) and prepares students for tertiary education. For the cohorts of study, the academic high school had two tracks: language and math. Vocational education is, in contrast, field-specific and prepares students for specific occupations; I group Information and Communication Technologies and Engineering (ISCED fields 61 and 71) as STEM.

Tertiary education (ISCED levels 5–8) consists of three types: vocational, professional, and academic. I refer to the latter two jointly as *college*. Similarly, I group vocational secondary and vocational tertiary educations as *vocational education*. A vocational secondary degree usually only gives direct access to vocational tertiary pro-

[38] They can also choose to enroll in an optional 10^{th} grade, which is a formal a continuation of primary school. In the analysis, I restrict the attention to enrollment in and completion of programs after primary school, i.e. after grade 9 and 10.

grams within the same specific field,[39] while an academic high school diploma gives access to all types of tertiary education. An application to tertiary education is an application to a specific program. Most college STEM programs require certain high school STEM courses as prerequisites, such as advanced Math and intermediate Physics and Chemistry. Therefore, an academic high school STEM diploma gives much easier access to college STEM majors than other secondary school degrees, although it is possible to take complementary courses after high school graduation. Acceptance to college mainly depends on high school GPA and most STEM programs admit all eligible applicants (or have very low GPA cutoffs).

To mirror the definition of field-specific STEM education to the one of STEM occupation, I define STEM in college as Physical Sciences, Mathematics, Statistics, Economics, Information and Communication Technologies, and Engineering (ISCED fields 53, 54, 311, 61, 71). However, the results are similar when including Biology. Another important reason for excluding biology is that women's underrepresentation in STEM is limited to math-intensive —and, generally, better paid—science fields (Kahn and Ginther, 2017). The analysis of STEM education considers field-specific STEM educations in any type and at any level of education after primary school. This is to not potentially confound the results on STEM choice with educational attainment. Thus, the main STEM outcomes of interest indicate whether the individual ever enrolls in and completes a field-specific STEM education preparing for the labor market, including secondary and tertiary vocational STEM programs and college STEM majors.

Moreover, I complement the main STEM measures with four additional outcomes; the results are reported in the appendix. I examine whether the first place of enrollment after primary school has a STEM focus, i.e. whether it is either secondary STEM vocational education or in the math track in the academic high school. In line with this, I consider the probability of ever completing the academic high school math track. Finally, I split field-specific STEM educations by type, thereby investigating effects on the probability of completing a vocational STEM program and a college STEM major, separately.[40]

A.4 Alternative Mechanisms

This appendix describes alternative mechanisms to the ones discussed in Subsection 5.1. These mechanisms cannot be the dominating ones, however, as they are not

[39]Students with a vocational secondary degree will often be required to have taken one or two academic high school courses at a basic level, such as Math and English. Many vocational secondary programs do not have a natural continuation at the tertiary level, though.

[40]Considering whether the highest completed education is within STEM reveals similar results as for having any field-specific STEM degree (not reported). Moreover, considering the probability of enrolling in the different types of STEM education rather than completing them also give similar results.

compatible with the empirical findings.

The effect of sibling interactions might also go in the opposite direction for two reasons. First, the spillover model in developmental psychology hypothesizes that siblings imitate and influence each other with their gender-specific traits. For instance, Brim (1958) and Koch (1955) show that mixed sex siblings exhibit more traits of the opposite gender and fewer of their own gender compared to same sex sibling pairs. Second, the reference group theory in sociology suggests that as soon as a same sex sibling is present in the family, the same sex sibling will be the child and parents' reference group (Butcher and Case, 1994). Therefore, having a same sex sibling might induce the child to behave more gender-stereotypically. Meanwhile, given the empirical findings, neither of these two theories can be the dominating mechanism for the effect of sibling gender composition on the development of women's gender identity.

Studies examining the relationship between sibling gender composition and educational attainment have argued that budget constraints may play an important role (Amin, 2009; Butcher and Case, 1994). If parents face no borrowing constraints, they should, according to standard economic theory, invest in each child until marginal costs equal marginal benefits. However, if parents face borrowing constraints, they might decide to allocate their financial resources depending on the gender composition of their children. If parents want income equality between their children and the returns to education are smaller for women than men, then having a brother instead of a sister would be beneficial. However, parental aversion to income inequality cannot be the dominating channel, as we would otherwise have observed that having a sibling of the opposite sex should make the educational choice less gender-stereotypical.

In contrast, parents might want to maximize the total income of their children, thereby investing more in the child with the greatest returns to education. If returns to education are larger for men than women, having a brother would have adverse effects on educational attainment. In support of this argument, Powell and Steelman (1989) find for students enrolled in one college in the U.S. that the number of brothers puts more pressure on parents' financial support than do the number of sisters. Nevertheless, this is not a likely mechanism in the Danish context because there is no tuition fee at any educational level. Moreover, students in vocational training typically receive apprenticeship wages and students in tertiary education receive governmental student grants and loans to cover living expenses. For all cohorts in the analysis, students in tertiary education have at least had access to a combination of grants and loans of 1,000 USD a month in 2017-prices. It is also less clear how borrowing constraints should affect field choice, given sibling gender composition has no effect on the probability of enrolling in any type of program after compulsory education. Moreover, a more recent study shows that, for later generations in the U.S., parents to at least one son compared to parents with no sons do not differentially invest in their daughters (Cools and Patacchini, 2017).

A.5 Appendix Tables and Figures

Table A5

Effect of Sibling Gender on STEM Education and Educational Performance

	STEM Focus in First Enrollment (1)	HS STEM Track Completion (2)	Vocational STEM Completion (3)	College STEM Completion (4)	Grade 9 Language GPA (5)	Grade 9 Math GPA (6)
Second-Born	-1.24***	-1.12***	-0.23**	-0.34***	-0.80	-0.80
Brother	(0.25)	(0.23)	(0.10)	(0.10)	(0.59)	(0.60)
Observations	103,769	103,769	103,769	103,769	82,978	82,350
Average	22.7	18.9	2.4	2.5	44.7	22.6

All estimates are multiplied by 100 to express effects in percentage points/percent of a standard deviation. Standard errors in parentheses; * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Main sample (first-born women born 1962–1975 with a second-born biological sibling born within four years apart) for STEM outcomes; girls born between 1986 and 1999 with the same selection criteria as for the main sample for the grade 9 outcomes. Each Column presents estimates from separate regressions. All models absorb fixed effects for birth municipality, year-by-month of birth, spacing in months to younger sibling, maternal age at birth, paternal age at birth, maternal level-by-field of education, and paternal level-by-field of education. The STEM outcome models further control for age at last observation in the education registry. *STEM Focus in First Enrollment* indicates whether the woman's first place of enrollment after compulsory schooling is in the academic high school math track or in a field-specific vocational STEM education. *HS STEM Track Completion* indicates whether the woman has completed the academic high school math track. *Vocational STEM Completion* indicates whether the woman has completed either secondary or tertiary vocational field-specific STEM education. *College STEM Completion* indicates whether the woman has completed a college degree or higher within STEM (excluding Biology). The Grade 9 GPA measures come from the written exam at the end of grade 9 in respectively Danish and Math and are standardized by year of graduation for the total population with mean zero and standard deviation of one.

Table A6

Principal Component Analysis: Parental Time Investment

	Mother		Father	
	Age 7	Age 11	Age 7	Age 11
First Principal Component				
Play	0.51	0.58	0.49	0.53
Homework	0.32	0.37	0.47	0.43
Out-of-school activity	0.39	0.45	0.38	0.51
Read/sing	0.49	0.40	0.47	0.34
Excursion	0.49	0.42	0.41	0.40
Eigenvalue				
First Component	1.54	1.63	1.81	1.84
Second Component	0.97	1.09	0.92	0.95

DALSC sample. Higher values reflect that parents do the specific activity more often.

Table A7

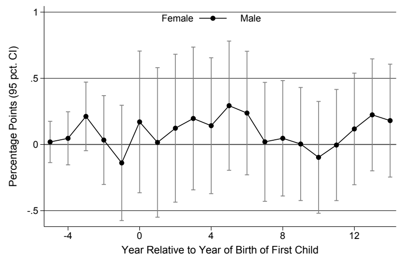
Effect of Sibling Gender on Components of Parental Time Investment at Age 7 and 11

	Play	Home-work	Out-of-School Activity	Read/ Sing	Excur-sion
	(1)	(2)	(3)	(4)	(5)
Panel A: Maternal Investment at age 7 ($N = 594$)					
Second-Born	0.10	0.09	0.01	0.04	0.14*
Brother	(0.08)	(0.08)	(0.09)	(0.08)	(0.08)
Panel B: Maternal Investment at age 11 ($N = 594$)					
Second-Born	0.15*	0.14	0.01	0.11	0.09
Brother	(0.09)	(0.09)	(0.09)	(0.09)	(0.09)
Panel C: Paternal Investment at age 7 ($N = 421$)					
Second-Born	-0.10	-0.19*	-0.01	-0.28***	-0.00
Brother	(0.10)	(0.10)	(0.10)	(0.10)	(0.10)
Panel D: Paternal Investment at age 11 ($N = 415$)					
Second-Born	-0.16	-0.22**	-0.11	-0.14	-0.08
Brother	(0.10)	(0.10)	(0.10)	(0.10)	(0.10)

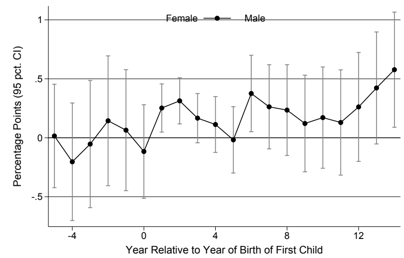
Standard errors in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. DALSC sample. Each Panel-Column represents the results from separate regressions. All models control for (quadratic) mother and father's age and fixed effects for spacing to the younger sibling in years, parental marital status in 1996, parents having been together for at least 5 years in 1996, region of birth, maternal level of education, paternal level of education, and family income level in 1995. Each of the individual components is standardized with mean zero and standard deviation of one.

Figure A1

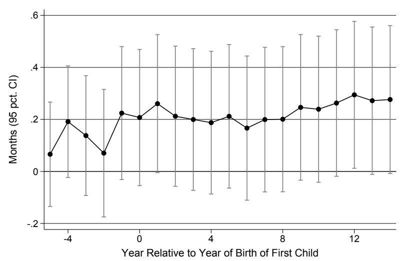
Parental Socio-Economic Status by Sibling Gender Composition



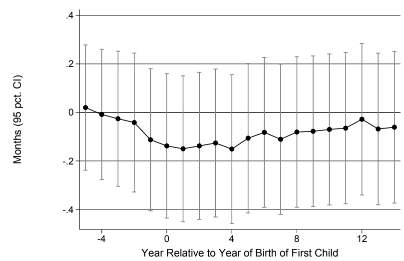
(a) Parents are Married



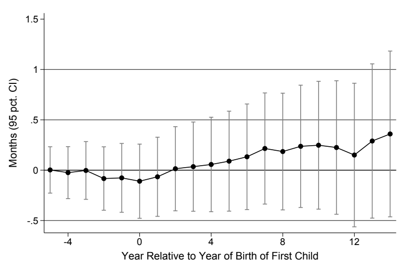
(b) Parents Cohabit/are Married



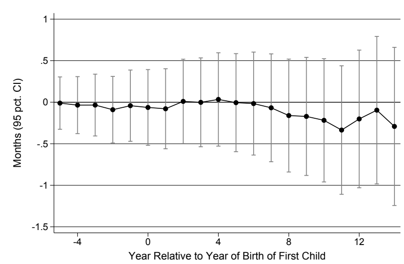
(c) Mother's Edu (months)



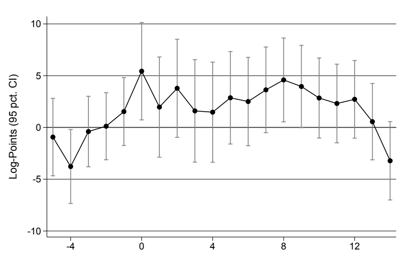
(d) Father's Edu (months)



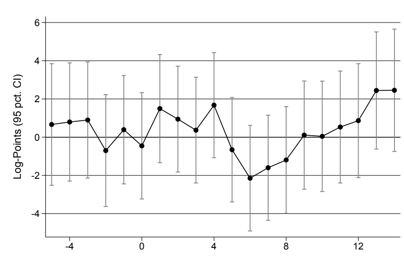
(e) Mother's Work Exp



(f) Father Work Exp



(g) Mother's log(Earnings)

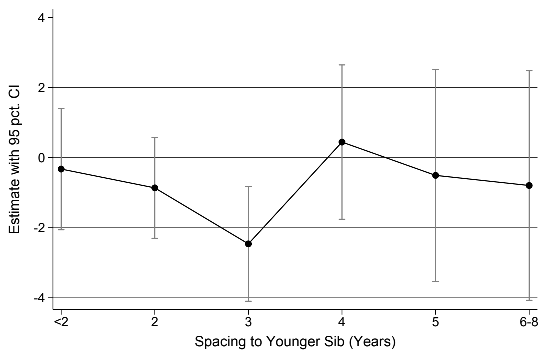


(h) Father's Log(Earnings)

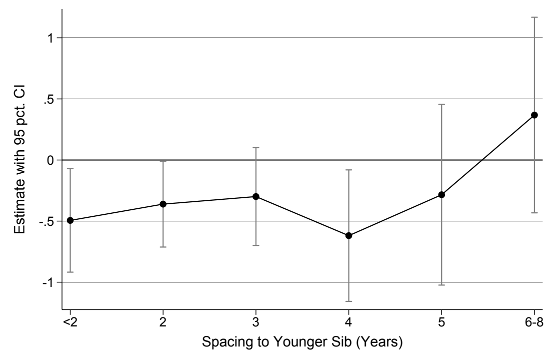
Note: Sample of first-born girls born between 1985 and 2002 with a second-born biological sibling born within four years apart. The whiskers represent the 95 percent confidence interval. All graphs illustrate the estimates from an event study of the effect of having a second-born brother. All models absorb time-specific fixed effects for birth municipality, year-by-month of birth, spacing in months to younger sibling, maternal age at birth, paternal age at birth, maternal level-by-field of education, and paternal level-by-field of education.

Figure A2

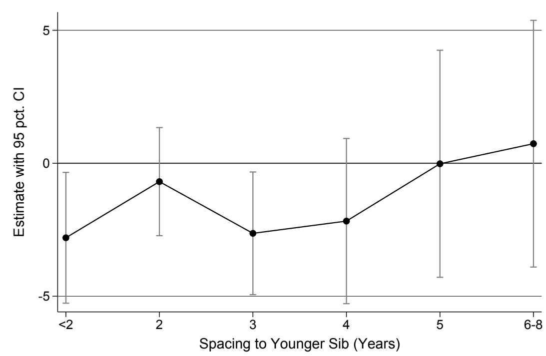
Effect of Sibling Gender on Choice of Occupation and Partner: Heterogeneity by Spacing



(a) Log(Male Share in Own Occupation)



(b) Share of Years in STEM Occupation

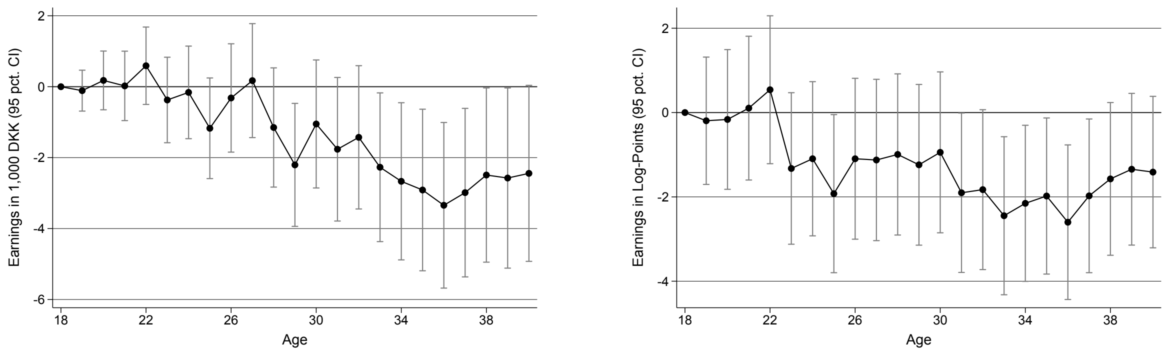


(c) Log(Female Share in Partner's Occupation)

Note: Main sample (first-born women born 1962–1975) including individuals with a second-born biological sibling born up to eight years apart. All graphs illustrate the estimated effect of having a second-born brother by birth spacing. The whiskers represent the 95 percent confidence interval. Each graph shows the estimates from a separate regression. All models absorb fixed effects for birth municipality, year-by-month of birth, spacing in months to younger sibling, maternal age at birth, paternal age at birth, maternal level-by-field of education, and paternal level-by-field of education. The models with own occupation also include dummies for the number of years observed in the income registry from age 31–40 and the number of years observed with a valid occupation code from age 31–40. For partner's occupation, the controls also include dummies for the partner's number of occupational observations and age at first and last observation. The occupational outcomes of the first-born women are measured as mean from age 31–40. The occupational outcome of the partner is measured mainly at ages 31–45 for the partner with whom the woman lived most years from age 31–41.

Figure A3

Effect of Sibling Gender on Earnings Age 18–40



(a) Labor Earnings **(b)** Log(Earnings)

Note: Main sample (first-born women born 1962–1975 with a younger biological sibling born within four years apart). The whiskers represent the 95 percent confidence interval. All graphs illustrate the estimates from an event study of the effect of having a second-born brother, where age 18 forms the base. Both models absorb time-specific fixed effects and individual fixed effects. *Labor Earnings* is measured in 1,000 DKK 2015-prices. *Log(Earnings)* is the natural logarithm of *Labor Earnings*.

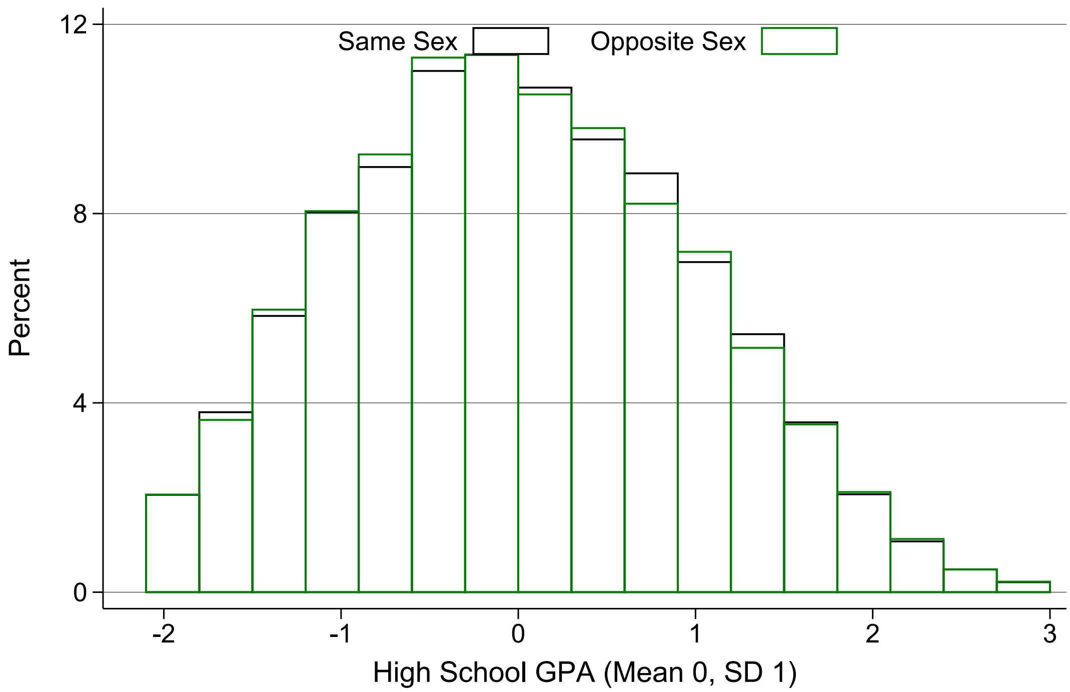
Table A8

Effect of Sibling Gender on Quality of Child-Parent and Child-Sibling Relations

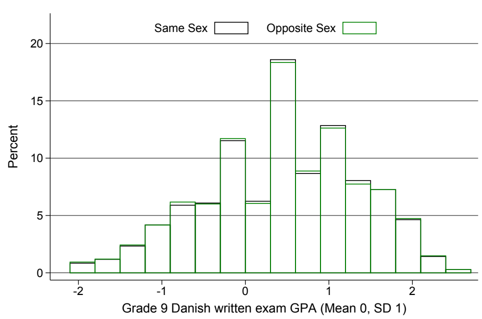
	Mother's	Father's	Child's relationship to		
	Relationship to Child		Mother	Father	Siblings
Child Age	11/15	7	15	15	15
	(1)	(2)	(3)	(4)	(5)
Second-Born	-0.12	-0.22**	0.07	-0.16*	-0.38***
Brother	(0.10)	(0.10)	(0.09)	(0.09)	(0.09)
Observations	441	434	498	489	485

Standard errors in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. DALSC sample. Each Column represents the results from separate regressions. All models control for (quadratic) mother and father's age and fixed effects for spacing to the younger sibling in years, parental marital status in 1996, parents having been together for at least 5 years in 1996, region of birth, maternal level of education, paternal level of education, and family income level in 1995. All child-parent relationship indexes represent the first component from principal component analyses, shown in Appendix Table A9, are standardized such that a higher value reflects a better relationship, the mean is zero, and the standard deviation is one. *Child's relationship to siblings* is an index of how easy the child thinks it is to talk to his/her siblings about matters that really bother her (standardized with mean zero and standard deviation of one).

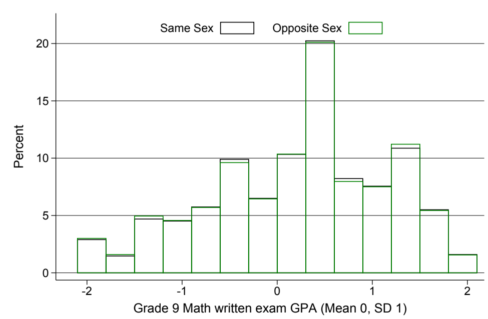
Figure A4
Distribution of Ability by Sibling Gender



(a) Academic High School GPA



(b) Grade 9 Language written exam



(c) Grade 9 Math written exam

Note: Main sample (first-born women born 1962–1975 with a second-born biological sibling born within four years apart) for academic high school GPA; girls born between 1986 and 1999 with the same selection criteria as for the main sample for the grade 9 outcomes. The Grade 9 GPA measures come from the written exam at the end of grade 9 in respectively Danish and Math. *Academic High School GPA* is observed for students completing the academic high school language and math tracks. The standardized GPA measures are standardized by year of graduation (for the high school GPA track-by-year of graduation) for the total population with mean zero and standard deviation of one. All graphs plot the distribution of the three measures of school performance by gender of the second-born sibling [sister (black) and brother (green)]. The tails are truncated to have at least five observations within each cell due to data protection rules.

Table A9
Principal Component Analysis: Child-Parent Relations

	Mother's	Father's	Child's rel'ship to	
	Rel'ship to Child		Mother	Father
First Principal Component				
Age 11: How close is the relationship between you and your daughter (1–4)?	0.71			
Age 15: How close is the relationship between you and your daughter (1–3)?	0.71			
Age 7: How close is the relationship between you and your daughter (1–4)?		0.71		
Age 7: Are you satisfied with the relationship between you and your daughter (1(yes)–2(no))?		0.71		
Age 15: Your mother/father plays a very big role in your life (1–5)			0.32	0.36
Age 15: Your relationship with your mother/father is important to you (1–5)			0.35	0.37
Age 15: Your mother/father loves you (1–5)			0.35	0.28
Age 15: You trust your mother/father (1–5)			0.38	0.40
Age 15: You can expect your mother/father to listen to you (1–5)			0.35	0.37
Age 15: You can go to your mother/father for advice (1–5)			0.40	0.36
Age 15: You can count on help from your mother/father if you have a problem (1–5)			0.36	0.37
Age 15: How easy is it to talk with your mother/father about matters that really bother you (1–5)			0.29	0.29
Eigenvalue				
First Component	1.34	1.25	4.07	4.53
Second Component	0.66	0.75	0.95	0.79

DALSC sample. All questions are answered on a likert scale with lower values being better. Therefore, the standardized measures used in Table A8 are all reversed, such that a higher value reflects a better relationship.

Table A10
Association Between First-Born Sibling's Gender and Second-Born Women's Gender Identity

	Log(Male Share in own Occ)	Works in STEM	Log(Female Share in Partner's Occ)	Log(Male Share in Edu)	Length (months)	High School GPA
	(1)	(2)	(3)	(4)	(5)	(6)
First-Born	-0.87*	-0.10	-1.20*	-1.15**	-0.42***	-0.05***
Brother	(0.46)	(0.11)	(0.67)	(0.52)	(0.15)	(0.01)
Observations	105,445	105,445	95,598	105,171	105,189	41,414
Average	787.6	4.623	292.2	333.6	156.9	0.002

Sample of second-born women born 1962–1975 with a first-born biological sibling born within four years apart. Each Column presents estimates from separate regressions. All models absorb fixed effects for birth municipality, year-by-month of birth, spacing in months to older sibling, maternal age at birth, paternal age at birth, maternal level-by-field of education, and paternal level-by-field of education. For the own occupation outcomes, basic controls also include dummies for the number of years observed in the income registry from age 31–40 and the number of years observed with a valid occupation code from age 31–40. For partner's occupation, basic controls also include dummies for the partner's number of occupational observations and age at first and last observation. The educational outcome models, except for high school GPA, further control for age at last observation in the education registry. The occupational outcomes of the second-born women are measured as mean from age 31–40. The occupational outcome of the partner is measured mainly at ages 31–45 for the partner with whom the woman lived most years from age 31–41. *Log(Male Share in Edu)* measures the natural logarithm of the share of men in the highest completed education (narrow field-by-level) by age 30. *Length* measures the length of the highest completed education in months by age 30. *HS GPA* measures final GPA from the academic high school and is standardized by track and year of graduation for the total population with mean zero and standard deviation of one.

Table A11

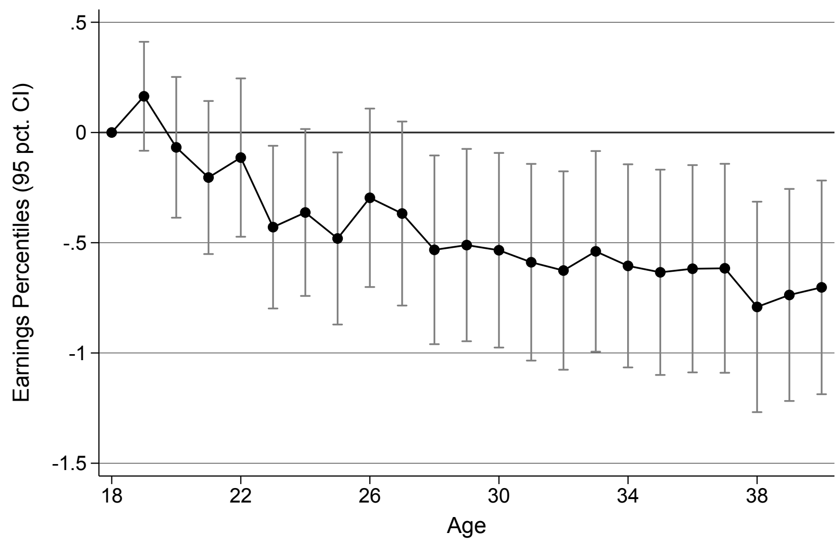
Men: Effect of Sibling Gender on Choice of Occupation

	(1)	(2)	(3)	(4)
Panel A: Log(Male Share in Own Occupation)				
Second-Born	0.51**	0.48*	0.44*	0.52**
Sister	(0.26)	(0.25)	(0.25)	(0.25)
Observations	108,366	108,365	108,365	108,365
Panel B: Share of Years Working in STEM Occupation				
Second-Born	0.44**	0.48***	0.51***	0.47***
Sister	(0.18)	(0.18)	(0.18)	(0.18)
Observations	108,366	108,365	108,365	108,365
Panel C: Share of Years Working as Manager				
Second-Born	-0.46***	-0.45***	-0.44***	-0.44***
Sister	(0.12)	(0.12)	(0.11)	(0.12)
Observations	108,366	108,365	108,365	108,365
No controls	✓			
Basic controls		✓	✓	✓
Parental education			✓	✓
Family size				✓

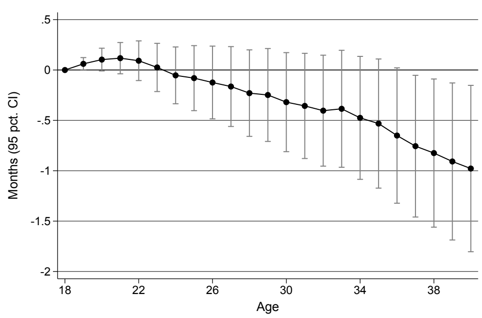
All estimates are multiplied by 100 to express effects in percentage/log-points. Standard errors in parentheses; * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Main sample (first-born *men* born 1962–1975 with a second-born biological sibling born within four years apart). Each Panel-Column presents estimates from separate regressions. *Basic controls* include fixed effects for birth municipality, year-by-month of birth, spacing in months to younger sibling, maternal age at birth, paternal age at birth, number of years observed in the income registry from age 31–40, and the number of years observed with a valid occupation code from age 31–40. *Parental education* controls include fixed effects for maternal level-by-field of education and paternal level-by-field of education. *Family size* controls include dummies for the number of biological siblings and dummies for the number of children the mother and father potentially have, respectively, from later relationships. The outcomes are measured as mean from age 31–40.

Figure A5

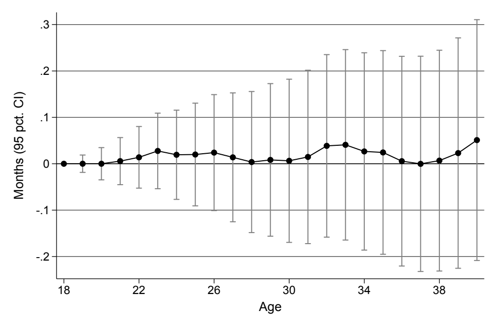
Men: Effect of Sibling Gender on Labor Market Outcomes Age 18–40



(a) Earnings Percentile



(b) Cumulated Work Experience



(c) Cumulated Unemployment

Note: Main sample (first-born *men* born 1962–1975 with a younger biological sibling born within four years apart). The whiskers represent the 95 percent confidence interval. All graphs illustrate the estimates from an event study of the effect of having a second-born brother, where age 18 forms the base. All models absorb time-specific fixed effects and individual fixed effects. *Earnings Percentile* measures the labor earnings percentile by age and cohort. *Work Experience* measures the cumulated lifetime work experience in months. *Unemployment* measures the cumulated lifetime unemployment in months.

Table A12

Men: Effect of Sibling Gender on Education and Family Formation

	(1)	(2)	(3)	(4)	(5)
Panel A: Education by age 30					
	Log(Male Share)	Length (months)	High School GPA	STEM Enrollment	STEM Completion
Second-Born Sister	0.47	-0.06	0.01	1.16***	0.45
	(0.29)	(0.15)	(0.01)	(0.30)	(0.28)
Observations	107,898	107,921	31,973	108,365	108,365
Panel B: Family Formation by age 41					
	Cohabit 18–41	Married 18–41	Has Any Children	# of Children	Age at First Birth
Second-Born Sister	-0.39***	-0.88***	-1.56***	-0.04***	0.09***
	(0.12)	(0.15)	(0.24)	(0.01)	(0.03)
Observations	108,365	108,365	108,365	108,365	86,124

Estimates in Columns (1), (4), and (5) in Panel A and Columns (1), (2), and (3) in Panel B are multiplied by 100 to express effects in percentage/log-points. Standard errors in parentheses; * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Main sample (first-born *men* born 1962–1975 with a second-born biological sibling born within four years apart). Each Panel-Column presents estimates from separate regressions. All models absorb fixed effects for birth municipality, year-by-month of birth, spacing in months to younger sibling, maternal age at birth, paternal age at birth, maternal level-by-field of education, and paternal level-by-field of education. The educational outcome models (except for high school GPA), further control for age at last observation in the education registry. *Log(Male Share)* measures the natural logarithm of the share of men in the highest completed education (narrow field-by-level) by age 30. *Length* measures the length of the highest completed education in months by age 30. *High School GPA* measures final GPA from the academic high school and is standardized by track and year of graduation for the total population with mean zero and standard deviation of one. *STEM Enrollment* indicates whether the man has ever enrolled in a field-specific STEM education at age 16–27. *STEM Completion* indicates whether the man has ever completed a field-specific STEM education by age 30. *Cohabit* measures the share of years age 18–41 during which the man has cohabited with a partner without being married. *Married* measures the share of years age 18–41 during which the man has been married. *Has Any Children* indicates whether the man has at least one child by age 41. *# of Children* measures the number of children the man has by age 41. *Age at First Childbirth* measures the age at the man's first childbirth in years, conditional on having any children.